



```
APPLICANT: Li, Xin-Liang
APPLICANT: Ljungdahl, Lars G.
APPLICANT: Chen, Huizhong
TITLE OF INVENTION: Lichenase and Coding Sequences
FILE REFERENCE: 55-96
CURRENT APPLICATION NUMBER: US/09/286,690
CURRENT FILING DATE: 1999-04-05
EARLIER APPLICATION NUMBER: US 60/027,882
EARLIER FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: PCT/US97/17811
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 8
LENGTH: 242
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-286-690-8
```

```
Query Match          17.1% Score 245.5; DB 3; Length 242;
Best Local Similarity 35.8%; Pred. No. 2,7e-17;
Matches 59; Conservative 24; Mismatches 71; Indels 11; Gaps 6;
```

```
QY 7 ECGAELTYLLEEVQYQKFEARKMAAAGTSSMFLYONGSEIADGRPVVEVDIEVLKRP 66
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 87 FCGENRSVQYGYGLYEVARKPAKNTGIVSSFFTYTGPTE---DGPMDIEDIEFLGKDT 143
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 67 GSFOSNIITGKAGAKTSEKHAAVSPAADAFTHTYGLETPNRYRMTVDGQEVKRTGGO 126
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 144 TVQGFNYVTNGG---NHKTIYDGFDAANAYHTAFDMQPSIKWYVGDG-LKHTATNO 199
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 127 VSNLTGTGCLRFNLWSSESA-AWVGCFDESKLPLEOFINWVYK 170
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 200 IPTTGGK--IMMNLNMGTCVDEMLGSYNGVN-PLVAHYMWVYTK 241
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```

## RESULT 3

```
US-08-103-998-2
Sequence 2, Application US/08103998
Patent No. 5470725
```

```
GENERAL INFORMATION:
APPLICANT: Borliss, Rainer
APPLICANT: Hofemeister, Jürgen
APPLICANT: Thomsen, Karl Kristian
APPLICANT: Olsen, Ole
APPLICANT: Von Wettstein, Dietrich
TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY & LARDNER
STREET: 1800 Diagonal Road, P.O. Box 299
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,998
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/773,652
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30307/123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
```

```
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-103-998-2
```

```
Query Match          17.0% Score 244; DB 1; Length 239;
Best Local Similarity 35.5%; Pred. No. 3,8e-17;
Matches 61; Conservative 26; Mismatches 69; Indels 16; Gaps 8;
```

```
QY 3 SAKDSGAELTYLLEEVQYQKFEARKMAAAGTSSMFLYONGSEIADGRPVVEVDIEVL 62
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 80 SYNKFDGGENRSVQYGYGLYEVARKPAKNTGIVSSFFTYTGPTE---GTPMDIEDIEFL 136
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 63 GRNPSFOSNIITGKAGAKTSEKHAAVSPAADAFTHTYGLETPNRYRMTVDGQEVKRT 122
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 137 GQDTTKVQGNVYTVNGVGH---EKVISLGFDAKGFHTYAFDMQPSIKWYVGDG-VLKHT 192
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 123 ECGVSNLTGTGCLRFNLWSSESA-AWVGCFDESKLPLEOFINWVYK 172
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 193 ---ATANIPSTPGKIMMNLNMGTCVDEMLGSYNGAN-PLVAYEDVY---KYT 237
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```

## RESULT 4

```
US-09-286-690-9
Sequence 9, Application US/09286690
Patent No. 6103511
```

```
GENERAL INFORMATION:
APPLICANT: Li, Xin-Liang
APPLICANT: Chen, Huizhong
TITLE OF INVENTION: Lichenase and Coding Sequences
FILE REFERENCE: 55-96
CURRENT APPLICATION NUMBER: US/09/286,690
CURRENT FILING DATE: 1999-04-05
EARLIER APPLICATION NUMBER: US 60/027,882
EARLIER FILING DATE: 1999-10-04
EARLIER APPLICATION NUMBER: PCT/US97/17811
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 9
LENGTH: 279
TYPE: PRT
ORGANISM: Clostridium thermocellum
US-09-286-690-9
```

```
Query Match          16.1% Score 231.5; DB 3; Length 279;
Best Local Similarity 35.3%; Pred. No. 9,1e-16;
Matches 61; Conservative 23; Mismatches 70; Indels 19; Gaps 7;
```

```
QY 5 KDFSGAELTYLLEEVQ-----YKFEARKMAAAGTSSMFLYONGSEIADGRPVVEVD 58
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 82 REYGSISYRKSEYRTKSEFGYGYEVARKAAKNYGVSSFFTYTGPTE---DNNWDEID 138
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 59 IEVLCKNPSFOSNIITGKAGAKTSEKHAAVSPAADAFTHTYGLETPNRYRMTVDGQOE 118
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 139 IEFLCKDTTKVQFNNYKKNVG---NEYLHNLGFDASQDFTYGGEMRPDYIDFVDSKK 195
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```

```
QY 119 VAKTEGQVSNLTGTGCLRFNLWSSESA-AWVGCFDESKLPLEOFINWVYK 169
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 196 VYR---GTRNIPVTPGKIMMNLNMGTCVDEMLGRYD-GRTPLAEXEYVYK 243
    |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```

## RESULT 5

```
US-09-286-690-10
Sequence 10, Application US/09286690
Patent No. 6103511
```

```
GENERAL INFORMATION:
```

```

; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-286-690-10
```

```
Query Match          15.9%; Score 229.5; DB 3; Length 243;
Best Local Similarity 33.7%; Pred. No. 1.2e-15;
Matches 57; Conservative 25; Mismatches 76; Indels 11; Gaps 6;
```

```
QY 3 SAKDSGAEIYTLAEVOYGFARMKMAASGVSSMFLYQNGSEIADGRPWEEVDIEVL 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 SYNKFPCGGENRSYQYTGGLYEVMKPAKNVGISSFFTYTGPT---DGPMDIEDIEFL 140
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 GKRRGSFOSNIITGKGAOKTSEKHHAVSPAADAQAFHTYGLWTPNVYRTVQGEVAKRT 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 141 GKRTTYVQEFYTYNGVG--NHEKIVNLGPDANSYHTYAFDQWPGIKWYVQD--LKHT 196
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 EGGQVSNLTGTGGLRPNLWSSESA-AWVGQFDESKLPLFOFIMWVKYK 170
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 197 ATQIQIQGTGK--IMNLMWAGVDEMIGSYN-GYTPPLSLHWRTK 242
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Bacillus polymyxa
US-09-286-690-7
```

```
Query Match          15.6%; Score 225; DB 3; Length 238;
Best Local Similarity 34.5%; Pred. No. 3.3e-15;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;
```

```
QY 7 FSGAEIYTLAEVOYGFARMKMAASGVSSMFLYQNGSEIADGRPWEEVDIEVLGNP 66
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 83 FDCEFRSTNNNGYGLYEVMKPAKNVGISSFFTYTGPSH---GTQWDEIDIEFLGKDT 139
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 67 GSFQSNIIITGKGAOKTSEKHHAVSPAADAQAFHTYGLWTPNVYRTVQGEVAKRTGEGQ 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 140 TKYQFNYYTYNGVGH--EKIINLGDASTSFTTYAFDQWPGIKWYVQD--VLKHT---A 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 127 VSNLTGTGGLRPNLWSSESA-AWVGQFDESKLPLFOFIMWVKYK 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
DB 193 TTNIPSPGKIMMNMWNGTGVDSMLGSYNGAN-PLVAEYDMV---KYT 236
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-286-690-11
```

```
Query Match          15.5%; Score 222.5; DB 3; Length 242;
Best Local Similarity 35.6%; Pred. No. 6.2e-15;
Matches 58; Conservative 22; Mismatches 64; Indels 19; Gaps 7;
```

```
QY 5 KDSFGAEIYTLAEVO-----YGFARMKMAASGVSSMFLYQNGSEIADGRPWEEVD 58
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 80 REYGGSYRSGEYRTKRSFPGYGLYEVMKPAKNVGISSFFTYTGPS--DNNPDEID 136
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 IEVLGKNPSSFOENIITGKGAOKTSEKHHAVSPAADAQAFHTYGLWTPNVYRTVQGEV 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 137 IEFLGKDTTYKVOFNWYKNGVG---NEXLHNLGPDASQDHTTYGFEMRPDYIDFYDGKK 193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 119 VRTEGGQVSNLTGTGGLRPNLWSSESA-AWVGQFDESKLPL 159
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 194 VYR----GTRNIVYTPGKIMMNMWNGTGVDSMLGGRD-GRTPL 231
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
; GENERAL INFORMATION:
; APPLICANT: Borrijs, Rainer
; APPLICANT: Hofemeister, Jurgen
; APPLICANT: Thomsen, Karl Kristian
; APPLICANT: Olsen, Ole
; APPLICANT: Von Wettstein, Dietrich
; TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLY & LARDNER
; STREET: 1800 Diagonal Road, P.O. Box 299
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,998
; FILING DATE:
; CLASSIFICATION: 435
```

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/773,652  
; FILING DATE: 16-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30307/123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-9300  
; TELEFAX: (703) 683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ. ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 237 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-103-998-4

Query Match 15.3%; Score 220.5; DB 1; Length 237;  
Best Local Similarity 33.9%; Pred. No. 9,6e-15;  
Matches 56; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

QY 7 FSGAEIYTLLEEVQYGFKFAKMAAASGVSMFLYONGSEIADGRPVEVDIEVIGKNP 66  
DB 82 FDCAEYRSTNYGGLYEVSMKPAKNGIVSFFTYGP--AHGTQWDEIDIEFLGKDT 138  
QY 67 GSFOSNITTKAGAKTSEKHNAVSPADAQAFHTYGLEMTPNYVRMTVDGQEVKTEGGO 126  
DB 139 TKVQFNYYTGAG---NHEKFADIGFPAANAYHTAFDMQPNSTIKMVDQO-LKHHTATTO 194  
QY 127 VSNLTGTGRLFNLMWSESA-AWVGQFDESKLPLFOFINVVKYK 170  
DB 195 IPAAPK--IMMILMNGTGVDMLGSYNGVN-PIYAHYDMRRYK 236

RESULT 9  
US-09-286-690-2  
; Sequence 2, Application US/09286690  
; Patent No. 6103511  
; GENERAL INFORMATION:  
; APPLICANT: Li, Xin-Liang  
; APPLICANT: Ljungdahl, Lars G.  
; APPLICANT: Chen, Huizhong  
; TITLE OF INVENTION: Lichenase and Coding Sequences  
; FILE REFERENCE: 55-96  
; CURRENT APPLICATION NUMBER: US/09/286,690  
; EARLIER FILING DATE: 1999-04-05  
; EARLIER APPLICATION NUMBER: US 60/027,882  
; EARLIER FILING DATE: 1999-10-04  
; EARLIER APPLICATION NUMBER: PCT/US97/17811  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ. ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ. ID NO: 2  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Orpiniomycetes sp. PC-2  
US-09-286-690-2

Query Match 13.1%; Score 189; DB 3; Length 245;  
Best Local Similarity 32.9%; Pred. No. 1.7e-11;  
Matches 56; Conservative 16; Mismatches 76; Indels 22; Gaps 8;

QY 5 KDFSGAEIYTLLEEVQ-----YGFKFAKMAAASGVSMFLYONGSEIADGRPVEVD 58  
DB 83 RDSGSG---YTCGSEYRTKNYGYGMFOVNMKPIKNPGVSSFFTYGPRS--DGTKWDEID 136  
QY 59 IEVLGNKPGSFOSNITTKAGAKTSEKHNAVSPADAQAFHTYGLEMTPNYVRMTVDGQ 118  
DB 137 IEFGLGIDTITVQGNYYTNGGHH---EHTIHYLFQDASQGHHTGFTPARASTIWIYVDGTA 193  
QY 119 VRTEGGVSNLTGTGQ-LRFNLMWSESA-AWVGQFDESKLPLFOFINV 166

DB 194 VYTA-----YDNIDPTGKIMMNAWNGIGVDWDLRPFN-GRTNISAYVDWV 238

RESULT 10  
US-08-824-707-2  
; Sequence 2, Application US/08824707  
; Patent No. 5919688  
; GENERAL INFORMATION:  
; APPLICANT: Ferrer, Pau  
; APPLICANT: Diers, Ivan  
; APPLICANT: Hedegard, Lisbeth  
; APPLICANT: Halkier, Torben  
; APPLICANT: Asenjo, Juan  
; APPLICANT: Savva, Demetris  
; TITLE OF INVENTION: No. 5919688el enzyme with beta-1,3-glucanase activity  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5919688o No. 5919688disk of No. 5919688ch America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10174-6401

COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,707  
; FILING DATE: 14-April-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gregg, Valera A.  
; REGISTRATION NUMBER: 35,127  
; REFERENCE/DOCKET NUMBER: 4290,204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ. ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 306 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-824-707-2

Query Match 10.7%; Score 154.5; DB 2; Length 306;  
Best Local Similarity 28.5%; Pred. No. 8.1e-08;  
Matches 53; Conservative 28; Mismatches 68; Indels 37; Gaps 11;

QY 7 FSGAEIYTLLEEV--QYGFKFAKMAAASGVSMFLYONGSEIADGRPV---EVDI-E 60  
DB 131 YTSARLTGQNVOPQGRIGRIARQIPROGIMSAFMV--GANLIP-TPMPTSGEIDNE 187  
QY 61 VLKKNP-----GSFOSNITTKAGAKTSEKHNAVSPADAQAFHTYGLEMTPNYVRVTGQ 116  
DB 188 NVGNAPHEVHGVTHGPGYSDNDGIMGTQHPQCSFADD--FHTPSIDMTPGSITLVDG 245  
QY 117 QEVKRTGEGVSNLTGTGRLFNLMWSESAWVGQFDESKLPLFOFINVVKYKYPGOG 176  
DB 246 QEVHRTVTDVG-----ANQWV--FDQ--PFLLLN-VALIGQWPGNP 283

QY 177 EGGSDF 182  
DB 284 DATPF 289  
RESULT 11  
US-09-159-106-2  
; Sequence 2, Application US/09159106  
; Patent No. 6284509



```

; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Halkier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; EARLIER FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRF
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-2

```

```

Query Match          10.1%; Score 145; DB 4; Length 263;
Best Local Similarity 25.9%; Pred. No. 6.1e-07;
Matches 49; Conservative 29; Mismatches 73; Indels 38; Gaps 10;

```

```

QY 19 QYKFFARKMAAASGVSMFLYQNGSELDGRPV---EVLGKNPFSQSNITG-- 76
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 82 QYGRLEARIOIPRGQ--IMPAFMMLGS--FPCTPMPSSGEIDIMENVG-FEPHRYHGV 136
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 77 -----KAGAKTSEKHHAASPAADQAFHTYGLEWTPNYRWTVDOEVRKTEGGVSNL 131
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 137 HGPYSGSGGIGTGMYPHOGMSFADTFHTFAVDKPGELTWVDDGQFHRVTRASVG--- 193
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 132 GTQGLRFNLMSSSAWVGQFDESKLPLFQFINWVKYKYPGQGGGS---DFTLDWT 187
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 194 -----ANAWV--FDQ---PFLLIN-VAVGQWMPGYPDGTQQLPQMKVVDY 234
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 188 DNFDFDGS 196
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 235 RYVDNGSGS 243

```

```

RESULT 12
US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Halkier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; EARLIER FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 435
; TYPE: PRF
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

```

```

Query Match          9.9%; Score 143; DB 4; Length 435;
Best Local Similarity 24.6%; Pred. No. 2.1e-06;
Matches 62; Conservative 29; Mismatches 95; Indels 66; Gaps 14;

```

```

QY 19 QYKFFARKMAAASGVSMFLYQNGSELDGRPV---EVDI-EVLGKNP-----GSRQ 70
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 134 QYGRLEARIOIPRGQ--IMPAFMMLGS--FPCTPMPSSGEIDIMENVGFEPRHRYHVA 190
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 71 SNITGKAGAKTSEKHHAASPAADQAFHTYGLEWTPNYRWTVDOEVRKTEGGVSNL 130
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 191 GPGYSGSGGI--TGMVQHOGMSFADTFHTFAVDKPGELTWVDDGQFHRVTRASVG--- 246
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 131 TGTQGLRFNLMSSSAWVGQFDESKLPLFQFINWVKYKYPGQGGGS---DFTLDW 186
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 247 -----ANAWV--FDQ---PFLLIN-VAVGQWMPGYPDGTQQLPQMKVVDY 286
    |||:||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 187 TDNDFDGS 196
    |||:||||:|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
DB 287 RYVDNGSGS 296

```

```

RESULT 13
US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Halkier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; EARLIER FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 303
; TYPE: PRF
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:14:27 ; Search time 12.6701 Seconds  
(without alignments)  
2025.857 Million cell updates/sec

Title: US-09-654-652a-2

Perfect score: 1439  
Sequence: 1 VSAKDFSGAEIYLTLEEVY.....PNSSVDKLAALLENHNNH 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1340	93.1	349	2 A44507	licheninase (EC 3.
2	266.5	18.5	851	2 H84053	endo-beta-1,3-1,4
3	265.5	18.5	252	2 A46378	licheninase (EC 3.
4	247.5	17.2	276	2 I40453	licheninase (EC 3.
5	245.5	17.1	242	1 LXBS	licheninase (EC 3.
6	242.5	16.9	802	2 A36910	xylinase, beta(1,3
7	238.5	16.6	239	1 A29091	licheninase (EC 3.
8	237	16.5	334	1 S23498	licheninase (EC 3.
9	229.5	15.9	243	1 S15388	licheninase (EC 3.
10	226	15.7	237	1 S11927	licheninase (EC 3.
11	225	15.6	238	1 S19012	licheninase (EC 3.
12	222.5	15.5	242	2 JS0611	endo-1,3(4)-beta-g
13	221	15.4	246	2 D97245	endo-1,3(4)-beta-g
14	206	14.3	263	2 A83055	endo-1,3-1,4-beta
15	206	14.3	289	2 C98231	endo-1,3-1,4-beta
16	197	13.7	269	2 H95976	endo-beta-1,3-1,4-
17	190.5	13.2	302	2 G84053	hybrid-endo-beta-1
18	182.5	12.7	507	2 S64507	probable membrane
19	180	12.5	268	2 S34804	exok protein - Rhi
20	172	12.0	282	2 T02354	xyloglucan endo-1,
21	171.5	11.9	277	2 B85354	hypothetical prote
22	169.5	11.8	282	2 A85354	hypothetical prote
23	169	11.7	642	2 B72428	lamnarinase - The
24	163	11.3	286	2 S71225	xyloglucan endo-1,
25	159	11.0	287	2 T04236	xyloglucan endo-1,
26	154.5	10.7	277	2 S71222	xyloglucan endo-1,
27	150.5	10.5	269	2 S61555	xyloglucan endo-1,
28	147	10.2	467	2 S30839	UTR protein - yea
29	146.5	10.2	284	2 T52097	xyloglucan endo-1,

30	142	9.9	422	2 S48564	probable membrane
31	141.5	9.8	289	2 T06166	xyloglucan endotra
32	140.5	9.8	280	2 T02090	xyloglucan endo-1,
33	140.5	9.8	310	2 A86239	protein T10024.17
34	140	9.7	305	2 G84568	probable xylogluca
35	139.5	9.7	286	2 T06202	xyloglucan endo-1,
36	133.5	9.3	286	2 S48201	licheninase (EC 3.
37	132	9.2	295	2 T10210	xyloglucan endo-1,
38	131.5	9.1	292	2 T06201	xyloglucan endo-1,
39	129.5	9.0	1324	2 T16265	endo-1,3(4)-beta-g
40	128	8.9	283	2 T07678	xyloglucan endo-1,
41	127.5	8.9	269	2 T05895	xyloglucan endo-1,
42	127	8.8	292	2 T04514	xyloglucan endo-1,
43	125.5	8.7	277	2 JF0156	end-xyloglucan tra
44	125	8.7	301	2 C87296	beta-glucanase [m
45	124	8.6	299	2 D84519	probable endoxylo

ALIGNMENTS

RESULT 1  
A44507  
licheninase (EC 3.2.1.73) - Fibrobacter succinogenes  
C:Species: Fibrobacter succinogenes  
C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 08-Oct-1999  
C:Accession: A44507  
R:Teacher, R.M.; Erfle, J.D.  
J. Bacteriol. 172, 3837-3841, 1990  
A:Title: DNA sequence of a Fibrobacter succinogenes mixed-linkage beta-glucanase (1,3  
A:Reference number: A44507; MUID:90299807; PMID:2193918  
A:Accession: A44507  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-349 <TEA>  
C:Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 93.1%; Score 1340; DB 2; Length 349;  
Best Local Similarity 99.6%; Pred. No. 1.5e-104;  
Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	VSAKDFSGAEIYLTLEEVYQGFARMKMAASGVSMFLYONGSEIADGRPVEVDIEY	61
DB	25	VSAKDFSGAEIYLTLEEVYQGFARMKMAASGVSMFLYONGSEIADGRPVEVDIEY	84
QY	62	LGNPFQSNITTKGAKAOKTSEKHNAAVSPADAQFHTYGLBMTPNYVWTVDQGEVRK	121
DB	85	LGNPFQSNITTKGAKAOKTSEKHNAAVSPADAQFHTYGLBMTPNYVWTVDQGEVRK	144
QY	122	TEGGVSNLTGTGJGLFENLMSSSAAMVGOFGDESKLPLFOFTIMVWVYVYKTPPGQGGGSD	181
DB	145	TEGGVSNLTGTGJGLFENLMSSSAAMVGOFGDESKLPLFOFTIMVWVYVYKTPPGQGGGSD	204
QY	182	FTLDWTDNFDTPFGSRMGKDMFTFDGNRVDLTDKNIYSRDSMLILALTRKGOESFNGOYP	241
DB	205	FTLDWTDNFDTPFGSRMGKDMFTFDGNRVDLTDKNIYSRDSMLILALTRKGOESFNGOYP	264
QY	242	RDDEPAPNSSS 252	
DB	265	RDDEPAPQSSS 275	

RESULT 2  
H84053  
endo-beta-1,3-1,4 glucanase (licheninase).bgls [Imported] - Bacillus halodurans (stra  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: H84053  
R:Takeuchi, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132



R:Murphy, N.; McConnell, D.J.; Cantwell, B.A.  
 Nucleic Acids Res. 12, 5355-5367, 1984  
 A:Title: The DNA sequence of the gene and genetic control sites for the excreted B. subtilis  
 A:Reference number: A93526; MUID:84272222; PMID:6087283  
 A:Accession: A22914  
 A:Molecule type: DNA  
 A:Residues: 1-203, 'L', 205-242 <MUR>  
 A:Cross-references: EMBL:X00754; NID:g93618; PIDN:CAA2328.1; PID:g685236  
 A:Experimental source: strain C120  
 A:Note: the authors translated the codon CAA for residue 29 as Lys and CCA for residue B  
 R:Tezuka, H.; Yuuki, T.; Yabuchi, S.  
 Agric. Biol. Chem. 53, 2335-2339, 1989  
 A:Title: Construction of a beta-glucanase hyperproducing *Bacillus subtilis* using the *clb*  
 A:Reference number: A90026  
 A:Accession: A90026  
 A:Molecule type: DNA  
 A:Residues: 1-23, 'S', 25-82, 'S', 84-242 <TEZ>  
 A:Cross-references: DDBJ:DD0518; NID:g216243; PIDN:BA00405.1; PID:g216244  
 A:Experimental source: strain Y-25, clone pLE100  
 R:Yuuki, T.; Tezuka, H.; Yabuchi, S.  
 Agric. Biol. Chem. 53, 2341-2346, 1989  
 A:Title: Purification and some properties of two enzymes from a beta-glucanase hyperprod  
 A:Reference number: A90027  
 A:Contents: annotation  
 A:Note: the amino ends of the mature forms of E-1 and E-2 are pyroglutamic acid and glut  
 C:Genetics:  
 A:Gene: *bgls*  
 C:Function:  
 A:Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and C  
 C:Superfamily: licheninase  
 C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation; F  
 F:1-28/Domain: signal sequence #status predicted <SIG>  
 F:29-242/Product: licheninase #status predicted <MAT>  
 F:29/Modified site: pyroglutamic carboxylic acid (Glu) (in mature form) (partial) #status

```

Query Match      16.9%; Score 242.5; DB 2; Length 802;-
Best Local Similarity 34.9%; Pred. No. 2.9e-12;
Matches    58; Conservative   21; Mismatches   74; Indels   13; Gaps   6;

QY      7 FSGAELTYLLEEVQVQKFEARKMKAASGVSSMFLYONGSELAIDGRPAVEVDIEFLGKNP 66
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      634 YSGGERITNNFYHYIYECSMOAMKNDGVSSFFYTGPS--DDNPWDEIDIELIGKNT 690
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      67 GSFSOINITGKAGAKORTSEKHNAVSPPADQAFTHTGLFEMTPENVRYRWVDOGEVRTEGGQ 126
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      691 TQVGNYNNTNGGKH---EKLVDLGFDSSSEAHYHGFPMQPNRYIAMUYDAGEHVRA----- 743
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      127 VSNLTGTGQ-LRFNLMSSEA-AWYGDFESKLPLEOFIMWKVKYK 170
       ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      744 TQDIKTPTGKIMMNAMPGITVDMDLKAFN-GTPTLTAHYOWVTYNK 788
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 7
A29091
licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens
N:Alternate names: 1,3-1,4-beta-D-glucan 4-glucanohydrolase; beta-glucanase; lichenass
C:Species: Bacillus amyloliquefaciens
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29091
R:Hofmeister, J.; Kurtz, A.; Borriess, R.; Knowles, J.
Gene 49, 177-187, 1986
A>Title: The beta-glucanase gene from Bacillus amyloliquefaciens shows extensive homo
A:Reference number: A91564; MUID:87192007; PMID:3106158
A:Accession: A29091
A:Molecule type: DNA
A:Residues: 1-239 <HOR>
A:Cross-references: GB:M15674; NID:g143009; PIDN:AAA8723.1; PID:g143010
A:Experimental source: strain BE20/78
C:Genetics:
A:Gene: bgla
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match      16.6%; Score 238.5; DB 1; Length 239;
Best Local Similarity 34.3%; Pred. No. 1.3e-12;
Matches    58; Conservative   26; Mismatches   74; Indels   11; Gaps   6;

QY      3 SAKDFSGAELYLLEIVQYQKFARMKMAASGVSSMFLYONGSEIADGRPWEVDIEVL 62
       |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      80 SYNKFCDCGENRSVQTGYGGLYEVRRMPAKMNGIVSSFYTGPRF--GTFPWDEIDIEFL 136
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      63 GKNPDSFOSNITGKAGAKORTSEKHNAVSPPADQAFTHTGLFEMTPENVRYRWVDOGEVRKT 122
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      137 GKDTTKRVQRNYNTNAG--NHKEFADLGFPDAAANYHYAPDAQWQNSIKRWVDDQ-LKRT 192
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY      123 ECGOVSNLTGTGGLSFENLMSSESA-AWYGDFESKLPLEOFIMWKVKYK 170
       |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      193 ATTQIPAPDGK--IMMNLMTGCVDDMLGSYNGVN-PITYAHDMWRKYRK 238
       |::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 8
S23498
licheninase (EC 3.2.1.73) liCB precursor - Clostridium thermocellum
N:Alternate names: beta-1,3-1,4-glucanase liCB; lichenase liCB
C:Species: Clostridium thermocellum
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23498; S22137
R:Schlummig, S.; Schwarz, W.H.; Staudenbauer, W.L.
Eur. J. Biochem. 204, 13-19, 1992
A>Title: Structure of the Clostridium thermocellum gene liCB and the encoded beta-1,3
  1 cellulases.
A:Reference number: S23498; MUID:92155194; PMID:1740123
A:Accession: S23498
A:Molecule type: DNA
A:Residues: 1-334 <SCH>
A:Cross-references: EMBL:X63355; NID:g40697; PIDN:CAAA4959.1; PID:g40698
C:Genetics:
A:Gene: liCB

```



```

Query Match 15.6%; Score 225; DB 1; Length 238;
Best Local Similarity 34.5%; Pred. No. 1.8e-11;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;

OY 7 FSGAELYLEEOYOGKFPARKMAAAGTGVSSMFLYONGSEIADGRPWVEVDIEVLGNP 66
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 FDCGEYRSTNNYGVGLYEVSMKPAKNTGIVSFFTYTGPSPH---GTQWDEIDIEFLGKDT 139
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 67 GSFOSNIITGKAGAKOAKTSEKHHAVSPAADOAFHYTGLEWTPYVRYWVDGCEVRKTEGGQ 126
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 TKVQFNYYTNGVGHH---EKIINIGFDASTSFHYAFEDMOGCIKWYDVG-VLKHT---A 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 127 VSNLFGTGG-LRPNLMSSESA-AWVGQFDESKLPLFGPINNVKRYKYN 172
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 TTNIPTSTGKIMNMLMNGTGVDSWLGSTYGAN-PLVAYEDWV---KYT 236
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
JS0611
endo-1,3(4)-beta-glucanase (EC 3.2.1.6) precursor - Clostridium thermocellum
N:Alternate names: laminarinase
C:Species: Clostridium thermocellum
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-1997
C:Accession: JS0611; S18726
R:Zverlov, V.V.; Laptev, D.A.; Tishkov, V.I.; Velikodvorskaia, G.A.
Biochem. Biophys. Res. Commun. 181, 507-512, 1991
A:Title: Nucleotide sequence of the Clostridium thermocellum laminarinase gene.
A:Reference number: JS0611; MUID:92095946; PMID:1755832
A:Accession: JS0611
A:Molecule type: DNA
A:Residues: 1-242 <ZVE1>
A:Cross-references: EMBL:X58392
R:Zverlov, V.V.; Velikodvorskaia, G.A.
Biotechnol. Lett. 12, 811-816, 1990
A:Title: Cloning the Clostridium thermocellum thermostable laminarinase gene in Escherichia coli.
A:Reference number: S18726
A:Accession: S18726
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-242 <ZVE2>
A:Cross-references: EMBL:X58392
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C:Genetics:
A:Gene: lam1
C:Superfamily: licheninase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-27/Domains: signal sequence #status predicted <sig>
F:28-242/Product: endo-1,3(4)-beta-glucanase #status predicted <END>

Query Match 15.5%; Score 222.5; DB 2; Length 242;
Best Local Similarity 35.6%; Pred. No. 3e-11;
Matches 58; Conservative 22; Mismatches 64; Indels 19; Gaps 7;

OY 5 KDFSGAELYTLEENG-----YGRFEARKMAAAGTGVSSMFLYONGSEIADGRPWVEVD 58
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 REYGGSYPRKSGEYTKSPFGYGEYVRKAKKNVGISSFYTYGPS--DNNPWDEID 136
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 59 IEVLAKNGSFPQSNITITKACAKOAKTSEKHHAVSPAADOAFHYTGLEWTPYVRYWVDQOE 118
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 IEFLLKDTTKQKQFMNKKNGVG---NEYLHNIGFPAASDFHYTGLEWTPYVRYWVDGKK 193
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 119 VRKTEGGVSNLTGTGG-LRPNLMSSESA-AWVGQFDESKLPL 159
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 VYR---GTRNIPVTPGKIMNMLMPGIGVDEWLGRYD-GRPPL 231
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
D97245
endo-1,3(4)-beta-glucanase family 16 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97245
R:Knilling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee

```

: Daly, M.J., Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: D97245  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-246 <KUR>  
A:Cross-references: GB:AEC01437; PIDN:AAK80751.1; PID:g15025848; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
.C.Genetics:  
A:Gene: CAC2807  
C:Superfamily: licheninase

Query Match 15.4%; Score 221; DB 2; Length 246;  
Best Local Similarity 33.7%; Pred. No. 4, le-11;  
Matches 58; Conservative 19; Mismatches 73; Indels 22; Gaps 8;

Oy 7 FSGAEIYLTLEEYOYGKFPEARKMAASGVSSMFLYONGSEINDGRPWVVDIEVGKNP 66  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::  
Db 91 YAGGEFRSNRRNYGYGLRYRSMKPRAKHIGVDSFEFTGSP--DNNPMWEIDIEFLEKDP 147  
67 GSFSNIITCKGAQNTSEKH---HAVSPAADAFFHTYGLEWTPNVVRMTVDDQEVRKTE 123  
Oy 148 TEVOFNYYTNGV-----KHELYIKLGFDAKGFFHYTGIMQNIAIMLVDSKEVYRA- 200  
124 GGVSNLGTQG-LRNPLWSESA-AWGQFDSEKLPLEOFINWVRVKYP 173  
Db 201 ---TSNIHPHGKVHMLWPGLGVDSWLGA VD-GVPVKAAYYWA--MYNP 245  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::

RESULT 14  
AB3055  
endo-1,3-1,4-beta-glycanase exok [imported] - Agrobacterium tumefaciens (strain C58,  
C.Species: Agrobacterium tumefaciens  
C.Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C.Accession: AB3055  
R.Wood, D.W.: Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Wo-  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCC-  
Science 294, 2317-2323, 2001  
A.Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka-  
ster, E.W.  
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AB3055  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-263 <KUR>  
A:Cross-references: GB:AEC08689; PIDN:AAL44856.1; PID:g17742502; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
.C.Genetics:  
A:Gene: exok  
A:Map position: linear chromosome  
C:Superfamily: licheninase

Query Match 14.3%; Score 206; DB 2; Length 263;  
Best Local Similarity 35.1%; Pred. No. 8e-10;  
Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

Oy 5 KPFSGAEIYTLTEEYOYGKFPEARKMAASGVSSMFLYONGSEINDGRPVEVDIEVLAK 64  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::  
Db 89 RNFAEIQTKGRKYRGTYEARMKATGSCLNSAFPTYIGPT--DKKPHDEIDFVELK 145  
65 NPGSPOSN-IITGKAGAQTSEKHAHVSPAADAFFHTYGLEWTPNVVRMTVDDQEVRKTE 123  
Db 146 NGKVQOLNDYIAAKG----NEKLPVEGGADGFMNDYAFVMEPORLRYYVNGLKLVH--- 198  
Oy 124 GGVSNLGT-----GTGQLRFNLMSSESAA-WVGP 152  
Db 199 --EVTDETKIIPONAKIFISLMCTDTLKDMXGF 230  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:14:27 ; Search time 11.7685 Seconds  
(without alignments)  
2025.857 Million cell updates/sec

Title: US-09-654-652A-1

Perfect score: 1333

Sequence: 1 MVSARDFSGAEALTYLEEVY.....TRKGQSFNGQVPRDEPAP 248

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1328	99.6	349	A44507	licheninase (EC 3.
2	266.5	20.0	851	H84053	endo-beta-1,3-1,4
3	265.5	19.9	252	A46378	licheninase (EC 3.
4	247.5	18.6	276	I40453	licheninase (EC 3.
5	245.5	18.4	242	1LXBS	licheninase (EC 3.
6	242.5	18.2	802	A36910	xylinase, beta(1,3
7	238.5	17.9	239	A29091	licheninase (EC 3.
8	237	17.8	334	S22498	licheninase (EC 3.
9	229.5	17.2	243	S15388	licheninase (EC 3.
10	226	17.0	237	S11927	licheninase (EC 3.
11	225	16.9	238	S19012	licheninase (EC 3.
12	222.5	16.7	242	J50611	endo-1,3(4)-beta-g
13	221	16.6	246	D97245	endo-1,3(4)-beta-g
14	206	15.5	263	A83055	endo-1,3-1,4-beta
15	206	15.5	289	C96231	endo-1,3-1,4-beta
16	197	14.8	269	H95976	endo-beta-1,3-1,4-
17	190.5	14.3	302	G84053	hybrid-endo-beta-1
18	182.5	13.7	507	S64507	probable membrane
19	180	13.5	268	S34804	exok protein - Rhl
20	172	12.9	282	T02354	xyloglucan endo-1,
21	171.5	12.9	277	B85354	hypothetical prote
22	169.5	12.7	282	B72428	hypothetical prote
23	169	12.7	642	B72428	laminarinase - The
24	163	12.2	286	S71225	xyloglucan endo-1,
25	159	11.9	287	T04236	xyloglucan endo-1,
26	154.5	11.6	277	S71222	xyloglucan endo-1,
27	150.5	11.3	269	S61555	xyloglucan endo-1,
28	147	11.0	467	S30839	UTR protein - yea
29	146.5	11.0	284	T52097	xyloglucan endo-1,

30	141.5	10.6	289	2	T06166	xyloglucan endotra
31	140.5	10.5	280	2	T02090	xyloglucan endo-1,
32	140.5	10.5	310	2	A86239	proteol T10024.17
33	140	10.5	305	2	G84568	probable xylogluca
34	139.5	10.4	286	2	T06202	xyloglucan endo-1,
35	138	10.4	422	2	S48564	probable membrane
36	133.5	10.0	285	2	S48201	licheninase (EC 3.
37	132	9.9	295	2	T10210	xyloglucan endo-1,
38	131.5	9.9	292	2	T06201	xyloglucan endo-1,
39	129.5	9.7	1324	2	T18265	endo-1,3(4)-beta-g
40	128	9.6	283	2	T07678	xyloglucan endo-1,
41	127.5	9.6	269	2	T05895	xyloglucan endo-1,
42	127	9.5	292	2	T04514	xyloglucan endo-1,
43	125.5	9.4	277	2	JF0156	end-xyloglucan tra
44	125	9.4	301	2	C87296	beta-glucanase [im
45	124	9.3	299	2	D84519	probable endoxylog

## ALIGNMENTS

## RESULT 1

A44507

licheninase (EC 3.2.1.73) - Fibrobacter succinogenes

C:Species: Fibrobacter succinogenes

C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 08-Oct-1999

C:Accession: A44507

R:Teacher, R.M.; Ertle, J.D.

J. Bacteriol. 172, 3837-3841, 1990

A:Title: DNA sequence of a Fibrobacter succinogenes mixed-1 linkage beta-glucanase (1,3

A:Reference number: A44507; MUID:90299807; PMID:2193918

A:Accession: A44507

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <TEA>

C:Cross-references: EMBL:M33676; NID:g148575; PIDN:AAA24896.1; PID:g148576

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 99.6%; Score 1328; DB 2; Length 349;

Best Local Similarity 100.0%; Pred. No. 1.5e-104;

Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2	VSAKDFSGAEALTYLEEVYQKGFARMKMAASGVSMFLYQNGSEIADGRPVEVDIEV	61
DB	25	VSAKDFSGAEALTYLEEVYQKGFARMKMAASGVSMFLYQNGSEIADGRPVEVDIEV	84
QY	62	LGNPQSFQSNITTKAGAQKTSKHHAVSPADAQFHYGLBMTPTNYKWFVDQGEVRK	121
DB	85	LGNPQSFQSNITTKAGAQKTSKHHAVSPADAQFHYGLBMTPTNYKWFVDQGEVRK	144
QY	122	TEGGVSNLTGTGGLRPNLWSSSAAMVGFDESKLPLOFIWVYVYKTPQGGGSGSD	181
DB	145	TEGGVSNLTGTGGLRPNLWSSSAAMVGFDESKLPLOFIWVYVYKTPQGGGSGSD	204
QY	182	FTLDWTDNFDPTGSRMGKDMFDGNRYDLTDKNITSRGMILALTRKGQSFNGQVP	241
DB	205	FTLDWTDNFDPTGSRMGKDMFDGNRYDLTDKNITSRGMILALTRKGQSFNGQVP	264
QY	242	RDDEPAP 248	
DB	265	RDDEPAP 271	

## RESULT 2

H84053

endo-beta-1,3-1,4 glucanase (licheninase) bgl5 [imported] - Bacillus halodurans (stra

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: H84053

R:Takekumi, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11056132



R.Murphy, N.; McConnell, D.J.; Cantwell, B.A.  
Nucleic Acids Res. 12, 5355-5367, 1984  
A>Title: The DNA sequence of the gene and genetic control sites for the excreted B. subtilis  
A:Reference number: A93526; MUID:84272222; PMID:6087283  
A:Accession: A22914  
A:Molecule type: DNA  
A:Residues: 1-203,'U',205-242 <MUR>  
A:Cross-references: EMBL:X00754; NID:g39818; PIDN:CAA25328.1; PID:g665236  
A:Experimental source: strain Cl120  
A>Note: the authors translated the codon CAA for residue 29 as Lys and CCA for residue 8  
R:Tezuka, H.; Yuki, T.; Yabuchi, S.  
Agric. Biol. Chem. 53, 2335-2339, 1989  
A>Title: Construction of a beta-glucanase hyperproducing Bacillus subtilis using the cld  
A:Reference number: A90026  
A:Accession: A90026  
A:Molecule type: DNA  
A:Residues: 1-23,'S',25-82,'S',84-242 <TEZ>  
A:Cross-references: DBD:DD0518; NID:g216243; PIDN:BA00405.1; PID:g216244  
A:Experimental source: strain Y-25, clone pLE100  
R:Yuki, T.; Tezuka, H.; Yabuchi, S.  
Agric. Biol. Chem. 53, 2341-2346, 1989  
A>Title: Purification and some properties of two enzymes from a beta-glucanase hyperprod  
A:Reference number: A90027  
A:Contents: annotation  
A>Note: source was hyperproducing strain HU-25 with gene from strain Y-25  
C:Genetics:  
A:Gene: bglIS  
C:Function:  
A>Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and c  
C:Superfamily: licheninase  
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation; F  
F:1-28/Domains: signal sequence #status predicted <SIG>  
F:29-242/Product: licheninase #status predicted <MT>  
F:29/Modified site: pyrrolic acid carboxylic acid (gin) (in mature form) (partial) #status

Query Match 18.4%; Score 245.5; DB: 1; Length 242;  
Best Local Similarity 35.8%; Pred. No. 2.2e-13;  
Matches 59; Conservative 24; Mismatches 71; Indels 11; Gaps 6;

Oy 7 ESGAEIYLEEVOYKFEARKMAAASGTSSMFLYONGSEIADGRPWVEDIEVLGNP 66  
| | : : : | : | | | | : | | | : | : | : | : | : |  
Db 87 FDGGEHSVQTYGAGXGLYEVRMKPAKNGTIVSSFFTYTGP--DTPDEIDIEFLGKDT 143  
| | : : : | : | | | | : | | | : | : | : | : | : |  
Oy 67 GSFGSLITIGKKAQCTSEHHAVSPAADAFFHTYGLEMTNRYRWTVDGCEVRKTGGQ 126  
| | : : : | : | | | | : | : | : | : | : | : | : |  
Db 144 TVQVFNYINGAG---NHKEIVDLGFDAANAHTYAFEDWPNQNSIKMYVDGQ-LKHATNQ 199  
| | : : : | : | | | | : | : | : | : | : | : | : |  
Oy 127 VSNLTGQGLRFNLMSSEA-AWVGDFESKLPTFOFINWKVKYK 170  
| | : : : | : | | | | : | : | : | : | : | : | : |  
Db 200 IFTTPCK--IMMLNMNGTGVDENLGSYNGVA-PLAHHDWRYYTK 241  
| | : : : | : | | | | : | : | : | : | : | : | : |

RESULT 6  
A36910  
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens  
C:Species: Ruminococcus flavefaciens  
C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 15-Oct-1999  
C:Accession: A36910  
R:Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.  
J. Bacteriol. 175, 2943-2951, 1993  
A>Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domai  
A:Reference number: A36910; MUID:9325938; PMID:8491715  
A:Accession: A36910  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-802 <FLI>  
A:Cross-references: GB:S61204; NID:g485910; PIDN:AAB26620.1; PID:g385911  
A>Note: sequence extracted from NCBI backbone (NCBIN\_131871, NCBIP\_131872)  
C:Superfamily: endo-1,4-beta-xylanase homology; Thermotoga xylanase A amino-terminal rep  
F:42-239/Domains: endo-1,4-beta-xylanase homology <YL>  
F:239-401/Domains: Thermotoga xylanase A amino-terminal repeat homology <TXA>

```
Query Match      18.2%, Score 242.5; DB 2; Length 802;  
Best Local Similarity 34.9%; Pred. No. 1.8e-12;  
Matches    58; Conservative   21; Mismatches   74; Indels     13; Gaps       6;
```

OY    7 FSGALTYLLEEVQCKFEARKMKMAAGCTVSMPFLXGSLFADGRPVEVDIEVLGNKP 66  
     :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db     YSGGEFRTNNFYHGYVCSQMAMKNDGVSSFFVTYGPS--DDNPWDEIDIELIGNKT 690

OY    67 GSFOENITTKACAGAOKTSECHHAHVSPADAOFHTYTGLEMTPENVRYATVGQEVRTEGGQ 126  
     ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     TQVCNPNNTYNNOGKH---EKLYDLGFDSSEAVHYTGPFMOQNITYAMYDGAEVRA----- 743

OY    127 VSNLTGTQG-LRFNIWSSEA-AWVGDFESKLPLFGFINWKYYK 170  
     ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     TODIKTPGKIIMNMAMPGLTVDDMLKAFN-GRPTLAHQWWTYNAK 788

RESULT 7  
A29091  
licheninase (EC 3.2.1.73) beta - Bacillus amyloliquefaciens  
N:Alternate names: 1.3-1,4-beta-D-glucan 4-gluconohydrolase; beta-glucanas; lichenas  
C:Species: Bacillus amyloliquefactions  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
R:Hofemeister, J.; Kurtz, A.; Borrijs, R.; Knowles, J.  
Gene 49, 177-187, 1986  
File Title: The beta-glucanas gene from Bacillus amyloliquefaciens shows extensive homo  
A:Reference number: A91504; MUID:87192007; PMID:3106158  
A:Accession: A29091  
A:Molecule type: DNA  
A:Residues: 1-239 <HOE>  
A:Cross-references: GB:M.5674; NID:g143009; PIDN:AAB87323.1; PID:g143010  
A:Experimental source: strain BE20/78  
C:Genetics:

A:Gene: bglA  
C:Superfamily: Licheninase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match              17.9% ; Score 238.5; DB 1; Length 239;  
Best Local Similarity    34.3%; Pred. No. 8.6e-13;  
Matches    58; Conservative    26; Mismatches    74; Indels    11; Gaps          6;

OY    3 SAKDPSGAELYTLBEVOYGEKFEPARMKMAASQTVMFLYONGSEIADGRPVEVDIEVL 62  
     |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     SYNKEFDGCENSVOYIGYLIVEVKRPAPKNNGIVSFETTYGPTE--GTFPWDEIDEIFL 136

OY    63 GKNPCGSFNSTTTGAGAOKTSEKHHAVSPAADOAFHTYGLFMTPNYRVKRWTDGOEVARKT 122  
     ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     GKDTTRKYVFNNYTNAG---NHKKFAIDGFPDAANAYHFADPDWPNSISKMYWDQO-LKHT 192

OY    123 ECGGVSNLTGTQGLRFNLMSSESAA-WMVGOPDESRLPLFGFINWKYYK 170  
     |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     ATTQLPAARGK--LKMNLMNGTCVDDWLGSTNGVN-PITYAHYDMWRRK 238

RESULT 8  
S23498  
licheninase (EC 3.2.1.73) liCb precursor - Clostridium thermocellum  
N:Alternate names: beta-1,3-1,4-glucanase liCb; Lichenase liCB  
C:Species: Clostridium thermocellum  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S23498; S22137  
R:Schlimmig, S.; Schwarz, W.H.; Straudenbauer, W.L.  
Eur. J. Biochem. 204, 13-19, 1992  
File Title: Structure of the Clostridium thermocellum gene liCb and the encoded beta-1,3  
li cellulases.

A:Reference number: S23498; MUID:92155194; PMID:1740123  
A:Accession: S23498  
A:Molecule type: DNA  
A:Residues: 1-534 <SCH>  
A:Cross-references: EMBL:X63355; NID:g40697; PIDN:CAAA44959.1; PID:g40698  
C:Genetics:  
A:Gene: liCb

C:Superfamily: licheninase l1cb: Clostridium cellulase repeat homology  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-334/Product: licheninase #status predicted <MMT>  
F:273-286/Domain: Clostridium cellulase repeat homology <CCR1>  
F:308-331/Domain: Clostridium cellulase repeat homology <CCR>

Query Match	17.8%	Score 237	DB 1	Length 334
Best Local Similarity	30.9%	Pred. No. 1.8e-12		
Matches 80	Conservative 27	Mismatches 94	Indels 58	Gaps 13

QY    5 KDFSALTYLLEVQ-----YGFELAKMMAASGVSSMFLIYNGSEIADGRPWVEVD 58  
      :: : : : : : : : : : : : : : : : : : :  
Db   82 REYGSGYPYKSGEYRTKSFPGYGYIEVMKAANKVGIVSSFYYIGPS---DNNPDEID 138

```

QY 59 IEVLGNPGRSQSNITGKAGAKTSEHHAVSADDAQFHTLTGLEWTPNVRWTVDGQE 118
      |||: | | : | | ||| | | : |||:
Db 139 IEF LGKDTTKVQFMWYKNGVGG--NEVLHNLGPDASQGFHTLTGGEWRPDYIDFVYDGKK 195

```

```

QY      119 VRKREGGVSILTGQG-LRFNLMSSEA-AWVGQFDESKLPLPFOFINWVAVY----- 169
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      196 VYR---GTRNIPVTPGKIMMNLMPGIGVDENLGRYD-GRTPLALEYEYVAVYYPNGVPQD 250

```

```

QY 170 -----KYP-----GGGE-GCSFET-----LDMTDNEDFDGDSRMGKG 201
      ||      |      |      |      |      |
Db 251 NPPTPTIAPSTPTINENLPLKGDVNGCGHVNSSDYSLFKRYLLRVIDRFVYGQDS---VA 307

```

```

qy      202 DWTFDGNRPVLTBDKNIIYSR 220
      |  | | | | | : |
db      308 DVNRDG-RIDSTDLTMLKR 325

```

RESULT 9  
S15388  
11chen1a000 (EC 3.2.1.73) - Bacillus licheniformis

N:Alternate names: endo-beta-1,3-1,4-D-glucanase, lichenase  
C:Species: Bacillus licheniformis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: G55200

**Reference number:** S5338  
**Article:** Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4  
**Author(s):** R. Lloberas, J. J. Perez-Pons, J. A. Querol, E. Eur. J. Biochem. 197, 337-343, 1991

A:Accession: S15388  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-243 <110>

A:Cross-references: EMBL:X51219; NID:g39558; PIDN:CAA0547.1; PID:g39559  
C:Superfamily: licheninase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query match	17.4%	Score 229.5	DB 1	Length 243
Best Local Similarity	33.7%	Pred. No. 5e-12		
Matches 57; Conservative	25;	Mismatches 76;	Indels 11;	Gaps

3 SAADFSGAELIILEEVYIGATLAAKMKAAASGVSSMTLQNSGLADGKRVVEVDIEVL 62  
84 SYNKFDGGEARSVQIYGYGLYLEVNMKPAKNVGISSFFIYTGPT--DGIPWDEIDIEFL 140

03 GANPGSFVSNLLTGAAGAAKTSERHHAVSPADQAFHLYGLEWIPNYVRMTVDGGEVKIT 127  
||: || :||| :| :| :| :|  
Db 141 GKDDTTKVGFNYYTNGVG---NHKEIVNLGFDAANSYHTYAEDMQPNSIKRYVDGQ-LKHT 196

123 EGGVSNLIGLIGLRLFLNWSSESA-AHWGQFDESKLPLEQFINWAKVVK 1/0  
 197 ATQGIQGTPECK--IMNNLNMGAGVDEPLGSYN-GVTPLSRSLHWRYTK 242

RESULT 10  
S11927  
phosphatase (EC 3.2.1.73) precursor [va] (dated) - Bacillus macerans

Malvaceae names: endo-beta-1,3-1,4-galactanase; lichenase  
C:Species: Bacillus macerans

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2000  
C:Accession: S11927  
R:Boiriss, R.; Buettner, K.; Meentsaetae, P.  
Mol. Gen. Genet. 222, 278-283, 1990  
A:Title: Structure of the beta-1,3-4-glucanase gene of *Bacillus macerans*: homolog  
A:Reference number: S11927; MUID:91109712; PMID:2274030  
A:Accession: S11927

A: Molecule type: DNA  
A: Residues: 1-237 <BOR>  
A: Cross-references: EMBL:X55959; NID:q296715; PIDN:CAA39426.1; PID:q296716

submitted to the Brookhaven Protein Data Bank, December 1994  
A: Reference number: A67074; PDB: 1MAC  
A: Contents: annotation: X-ray crystallography, 2.3 angstroms, residues 26-37, 'P', 39-2

R. Kretzel, T. Simon, O. Borriess, R. Hennemann, U. Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993

A. Title: Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase.

A:Contents: annotation: X-ray crystallography, 2.3 angstroms  
A.Note: recombinant form with residues 26-41 from *Bacillus amyloliquefaciens* expressed  
C:Function:

C:Superfamily: licheninase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F1-25/Domain: signal sequence #status predicted <Pro>

E-55-84/Disulfide bonds: #status experimental  
E-128/Active site: Glu #status predicted

Best Local Similarity 35.1%; Pred. No. 9.7e-12;  
Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

Db 82 FDCAEYRSTNIGYGLIEVSMKPAKNTGIVSSFFYYTGP--AHGTQWDEIDIEFLGKDT 138

```

Db      139  TKVDPNNYYINGVCGH-----EKVISTSLGFDASKGCHTAAFDMQPGIIKMWYDVG-VLKHH---A 191

```

```

Db      192 TANISTPQKIMNMLNMGTCYDDNMLGSLYNGAN-PLXAEIDWV---KTY 235
          ::| | | : |||: | : | :: : ||: |||

```

RESULT 11  
SI9012  
licheninase (EC 3.2.1.73) precursor - *Bacillus polymyxa*

C:Species: Bacillus polymyxa  
C:Date: 23-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 18-Jun-1999  
C:Accession: S19012

J. Bacteriol. 173, 7705-7710, 1991  
A:Title: Two beta-glycanase genes are clustered in *Bacillus polymyxa*: molecular cloning  
A:Reference number: S15011; MUID:92041687; PMID:1938968

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <GOS>

```

C:Function: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and  
A:Description: licheninase  
C:Superfamily: licheninase

F:27-237/Product: licheninase #status predicted <MAT>  
E:56-85/disulfide bonds: #status predicted  
F:129/Active site: Glu #status predicted

```

Query Match      16.9%; Score 225; DB 1; Length 238;
Best Local Similarity 34.5%; Pred. No. 1.2e-11;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;

OY    7 FSGAELTYLLEEVGYGKFEARKMKMAASGVTSMPFLYONGSEIADGRPWVEVIEVLGNP 66
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
DB    83 FDGGEYRSTNNNGCYGLYEYSKPAKNKGIVSFFLYTGPSPH---GTOMDEIDIEFLGKDT 139
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
OY    67 GSFOSNITIGKAGAOKTSEKHAAVSPPADOAFHTYGLEMTPNYRWTVDGOEVRKTEGQ 126
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
DB    140 TKQOFNYVTGVCGH---EKINLGFDASTSFHTYAFPMDPCYIKWYVDG-VLKHT---A 192
       :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
OY    127 VSNLTGTQG-LRFNLMWSSEA-AWVGOFDESKLPFQFINWVKYKYKT 172
       :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
DB    193 TTNIPSTPGKIMMNLWGTVGSWLGSYGAN-PLYAEYDWY---KYT 236
       :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

RESULT 12
J50611
endo-1,3(4)-beta-glucanase (EC 3.2.1.6) precursor - Clostridium thermocellum
N:Alternate names: laminarinase
C:Species: Clostridium thermocellum
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-1997
A:Accession: J50611.s18726
R:Zverlov, V.V.; Laptev, D.A.; Tishkov, V.I.; Velikodvorskaja, G.A.
Biochem. Biophys. Res. Commun. 181, 507-512, 1991
A>Title: Nucleotide sequence of the Clostridium thermocellum laminarinase gene.
A:Reference number: J50611; MID:92095946; PMID:1755832
A:Accession: J50611
A:Molecule type: DNA
A:Residues: 1-242 <ZVE1>
A:Cross-references: EMBL:X58392
R:Zverlov, V.V.; Velikodvorskaya, G.A.
Biotechnol. Lett. 12, 811-816, 1990
A>Title: Cloning the Clostridium thermocellum thermostable laminarinase gene in Escherichia coli.
A:Reference number: S18726
A:Accession: S18726
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-242 <ZVE2>
A:Cross-references: EMBL:X58392
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1991
C:Genetics:
A:Gene: Lam1
C:Superfamily: Licheninase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-27/Domains: signal sequence #status predicted <SIG>
F:28-242/Product: endo-1,3(4)-beta-glucanase #status predicted <END>

Query Match      16.7%; Score 222.5; DB 2; Length 242;
Best Local Similarity 35.6%; Pred. No. 2e-11;
Matches 58; Conservative 22; Mismatches 64; Indels 19; Gaps 7;

OY    5 KDKSSGAELYLLEVQ-----YCKFEARKMKMAASGVTSMSMLYONGSEIADGRPVEVD 58
       :   :   :   |   |   |   |   |   |   |   |   |   |   |   |   |
DB    80 REFGGSPIPKSYGSEFRKRSFEGYGEYVRMKAKNVGISSEFTYGPS--DNPPWEID 136
       :   :   :   |   |   |   |   |   |   |   |   |   |   |   |   |
OY    59 IEVLGNPDSFOSNIITIGKAGAOKTSEKHAAVSPPADOAFHTYGLEMTPNYRWTVDGOE 118
       |||||   |   |   |   |   |   |   |   |   |   |   |   |   |
DB    137 IEFLLGDTRTKVQNMWKYNGVG---NEYLNHLGFDASDFHTYGFGEEMRPDIIDFYDGKK 193
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
OY    119 VRKTEGGOVSNLTGTQG-LRFNLMWSSEA-AWVGOFDESKLP 159
       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
DB    194 VYR----GTFNIPVTPGKIMMNLWPICIGVDMLGRYD-GRPL 231
       :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

RESULT 13
D97245
endo-1,3(4)-beta-glucanase family 16 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97245
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
```

Y. Daly, M.J., Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96500; MUID:21359325; PMID:21359325  
A:Accession: D97245  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-246 <KUR>  
A:Cross-references: GB:A001437; PIDN:AAK80751.1; PID:g15025848; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2807  
C:Superfamily: licheninase

Query Match 16.6%; Score 221; DB 2; Length 246;  
Best Local Similarity 33.7%; Pred. No. 2, 7e-11;  
Matches 58; Conservative 19; Mismatches 73; Indels 22; Gaps 8;

OY 7 FSGAEIYLTEEVOYGKEFARMMAASGVSMFLYONGSEIADGRPWVEVDIEVLGKNP 66  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 91 YAGGERSNRRXCYGLVPSMKPAKHIGVDSSEFFSYTGPS--DNNPMDEIDIEFLGKDT 147  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 67 GSFGNIITGKAKAQTSKHK---HAVSPAADAQAFHTTGLEMTPNVRVTVDGQEVKRKE 123  
Db 148 TEVGNIYYNYNGV-----KHEYLYKLGFDAKSGFTHTYIMEONYIAMLVDSKEYRYRA- 200  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 124 GGQVENLTGTCT-LRFNLWSSESA-AWGQPFESKRLPFOFIWVYVYKYTP 173  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 201 ---TSNIPRHPRKVAMNLMPPGIGVDSMLGAYD-GVTPVAAYIYMA---MYNP 245  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 14  
AB3055  
endo-1,3-1,4-beta-glycanase exok [imported] - Agrobacterium tumefaciens (strain C58,  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C:Accession: AB3055  
R.Wood, D.W.: Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Wo-  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCC-  
r. Kapf, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Ka-  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AB3055  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-263 <KUR>  
A:Cross-references: GB:AE008689; PIDN:AAL44856.1; PID:g17742502; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: exok  
A:Map position: linear chromosome  
C:Superfamily: licheninase

Query Match 15.5%; Score 206; DB 2; Length 263;  
Best Local Similarity 35.1%; Pred. No. 5, 4e-10;  
Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

OY 5 KDEGSAELYTEEVOYGKEFARMMAASGVSMFLYONGSEIADGRPWVEVDIEVLGK 64  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 89 RNFAEGEIQTKRGYRYGEARMKATGSLNSAFFTYIGPT---DKKHDEIDFEVLGK 145  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 65 NNGSQSN-IITGKAKAQTSKHNHNAVSPAADAQAFHTTGLEMTPNVRVTVDGQEVKRKE 123  
Db 146 NGKGQVLNDIYAAGG---NEKLVPVEGAAGAFNDYAFVMEPORLRYVYNGKLVH--- 198  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
OY 124 GGQVENLTGTCT-LRFNLWSSESA-AWGQPFESKRLPFOFIWVYVYKYTP 173  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 199 ---EVTDETIRPONAKIFFSINGTDLTKDMCKF 230  
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

## RESULT 15

C98231

endo-1,3-1,4-beta-glucanase exok precursor (succinoglycan biosynthesis protein exok) [in

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence: revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: C98231

R:Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:11743194

A:Accession: C98231

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 &lt;KUR&gt;

A:Cross-references: GB:AE007870; PIDN:AAK89373.1; PID:q15159224; GSPDB:GN00170

C:Genetics:

A:Gene: AGR\_L\_1600

A:Map position: linear chromosome

C:Superfamily: licheninase

Query Match 15.5%; Score 206; DB 2; Length 289;  
Best Local Similarity 35.1%; Pred. No. 6e-10; Mismatches 56; Indels 18; Gaps 6;

Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

QY 5 KDFSGAELTYLLEVOYGFARMAAASQVSMFLYONGSEIADGRPMVEVDIEYLGR 64

DB 115 RNFACEIQTGKRYRYGTYEARKAKATGSLNSAFFTYIGPT---DKPHDEIDFEYLGR 171

QY 65 NPGSFQSN-ITGKAGAQKTSEKHNHVASPAADQAFHTYGLEWTPNYVRWTVDSGQEVYRTE 123

DB 172 NTGKYQLNQYIAAKGC---NEKLVPEEGADGFNDYAFVMEPQRLRYVNGKLVH--- 224

QY 124 GGOVSNLT---CTOGLRFNLMSSESAA-WVQGF 152

DB 225 --EVTDETKIPQNAQKIFPSLWGTDTLKDWMGKF 256

QY 124 GGOVSNLT---CTOGLRFNLMSSESAA-WVQGF 152

DB 225 --EVTDETKIPQNAQKIFPSLWGTDTLKDWMGKF 256

Search completed: January 9, 2003, 12:18:58

Job time : 12.7685 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:49:36 : Search time 13.7338 Seconds  
(without alignments)  
747.688 Million cell updates/sec

Title: US-09-654-652A-3  
1824  
Sequence: 1 MNIKRTAVKSAALAVAAAAA.....AKGAKVNPNGHKRYRVAPEH 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents-AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCtUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210	66.3	228	3	US-09-286-690-12
2	248	13.6	239	1	US-08-103-998-2
3	248	13.6	242	3	US-09-286-690-8
4	233.5	12.8	243	3	US-09-286-690-10
5	231.5	12.7	279	3	US-09-286-690-9
6	227.5	12.5	238	3	US-09-286-690-7
7	224.5	12.3	242	3	US-09-286-690-11
8	222.5	12.2	237	1	US-08-103-998-4
9	190.5	10.4	245	3	US-09-286-690-2
10	159.5	8.7	205	2	US-08-824-707-2
11	145	7.9	263	4	US-09-159-106-2
12	144	7.9	435	4	US-09-159-106-11
13	141.5	7.8	303	4	US-09-159-106-13
14	133.5	7.3	326	2	US-08-712-072C-4
15	119	6.5	321	2	US-08-712-072C-3
16	117.5	6.4	289	2	US-08-580-545B-4
17	117.5	6.4	289	4	US-09-262-653A-4
18	117.5	6.4	289	4	US-08-867-484A-2
19	117	6.4	620	4	US-09-269-731-6
20	116.5	6.4	654	1	US-08-392-828C-2
21	116.5	6.4	654	1	US-09-330-945-2
22	115	6.3	262	1	US-08-392-828C-37
23	115	6.3	262	3	US-09-330-945-37
24	114.5	6.3	545	4	US-09-269-731-8
25	113	6.2	285	2	US-08-712-072C-5
26	110.5	6.1	295	1	US-08-295-657-3
27	110	6.0	287	4	US-08-640-737-8

28	109.5	6.0	289	4	US-08-640-737-2	Sequence 2, Appli
29	109	6.0	382	4	US-08-213-419B-19	Sequence 19, Appli
30	108.5	5.9	367	4	US-08-213-419B-15	Sequence 15, Appli
31	108.5	5.9	989	4	US-08-213-419B-2	Sequence 2, Appli
32	108.5	5.9	983	4	US-09-213-419B-4	Sequence 4, Appli
33	108	5.9	583	4	US-09-620-412C-353	Sequence 353, App
34	107.5	5.9	666	4	US-08-961-083-2	Sequence 2, Appli
35	106.5	5.8	582	3	US-08-481-435-6	Sequence 6, Appli
36	105	5.8	422	2	US-08-712-072C-2	Sequence 2, Appli
37	104	5.7	348	4	US-09-216-295-16	Sequence 16, Appli
38	103.5	5.7	395	2	US-08-404-531B-3	Sequence 3, Appli
39	103.5	5.7	395	3	US-08-476-900A-3	Sequence 3, Appli
40	103.5	5.7	395	3	US-08-488-546A-3	Sequence 3, Appli
41	103.5	5.7	1580	4	US-08-726-320-1	Sequence 1, Appli
42	103.5	5.7	1580	4	US-09-208-716-1	Sequence 1, Appli
43	103.5	5.7	1581	4	US-08-726-320-3	Sequence 3, Appli
44	103.5	5.7	1581	4	US-09-208-716-3	Sequence 3, Appli
45	98.5	5.4	1581	2	US-08-404-531B-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Hui-zhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match          66.3%; Score 1210; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 1,3e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNIKRTAVKSAALAVAAAAALTTNWSAKDFSGAELTYLEVOYGFPEARMKMAASGTVS 60
DB 1 MNIKRTAVKSAALAVAAAAALTTNWSAKDFSGAELTYLEVOYGFPEARMKMAASGTVS 60
QY 61 SMFLVONGSEIINDGRPWVEVDIEVLGKNPQSFQSNITTKAKAOKTSEKHNHVAAPDA 120
DB 61 SMFLVONGSEIINDGRPWVEVDIEVLGKNPQSFQSNITTKAKAOKTSEKHNHVAAPDA 120
QY 121 PFTYGLEWPNVYRMTVDQOEVRKTEGGVSNLTGQGRFLMSESAAMVGOPEDESKL 180
DB 121 PFTYGLEWPNVYRMTVDQOEVRKTEGGVSNLTGQGRFLMSESAAMVGOPEDESKL 180
QY 181 PLFOFLNMVKKVYKTPPGQEGGSDFTLMDTNDFFDGSRMCKGDMTF 228
DB 181 PLFOFLNMVKKVYKTPPGQEGGSDFTLMDTNDFFDGSRMCKGDMTF 228

RESULT 2
US-08-103-998-2
; Sequence 2, Application US/08103998
; Patent No. 5470725
; GENERAL INFORMATION:
```

```

: APPLICANT: Borris, Rainer
: APPLICANT: Hofemeister, Jürgen
: APPLICANT: Thomsen, Karl Kristian
: APPLICANT: Olsen, Ole
: APPLICANT: Von Wettstein, Dietrich
: TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY & LARDNER
: STREET: 1800 Diagonal Road, P.O. Box 299
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/103,998
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/773,652
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Bent, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30307/123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-9300
: TELEFAX: (703) 683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 239 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-103-998-2

Query Match 13.6%; Score 248; DB 1; Length 239;
Best Local Similarity 35.4%; Pred. No. 6.8e-15;
Matches 62; Conservative 27; Mismatches 70; Indels 16; Gaps 8;

Qy 23 TNVSAKDFSGALVLTLEEVYQGFPEARKMAASGVSMFLYONGSEIADGRPWVEVDI 82
Db 77 TSPSYNKFDCGEMRSVQYTGYLEVRMKPAKNTGIVSFFTYGTPT--GTPMDEIDI 133

Qy 83 EYLGNKPGSFQSNITITGKAGAKTSEKHHAVSPADQAFHTYTGLEMTPNYVRYRWYDGOEV 142
Db 134 EFLGKDTTKVQFNYYTNGVGH--EKVYISLGFDAASKGHTYAFDMQPCYIMYVDG-VL 189

Qy 143 RKTGGVSNLTGTGQ-LRFNLMSSEA-AWVGQFDESKLPLEQFINWYKVKYK 193
Db 190 KHT--ATAINPSTPGKIMMNLNMGTVDMWLGSYNGAN-PLYAEVDWV---KYT 237

RESULT 3
US-09-286-690-8
: Sequence 8, Application US/09286690
: Patent No. 6103511
: GENERAL INFORMATION:
: APPLICANT: Li, Xin-Liang
: APPLICANT: Ljungdahl, Lars G.
: TITLE OF INVENTION: Lichenase and Coding Sequences
: FILE REFERENCE: 55-96
: CURRENT APPLICATION NUMBER: US/09/286,690
: CURRENT FILING DATE: 1999-04-05
: EARLIER APPLICATION NUMBER: US 60/027,882
: EARLIER FILING DATE: 1999-10-04
```

```

: EARLIER APPLICATION NUMBER: PCT/US97/17811
: EARLIER FILING DATE: 1997-10-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 8
: LENGTH: 242
: TYPE: PRT
: ORGANISM: Bacillus subtilis
: US-09-286-690-8

Query Match 13.6%; Score 248; DB 3; Length 242;
Best Local Similarity 36.0%; Pred. No. 6.9e-15;
Matches 63; Conservative 25; Mismatches 75; Indels 12; Gaps 7;

Qy 20 ALTTNVAKDFSGALVLTLEEVYQGFPEARKMAASGVSMFLYONGSEIADGRPWVE 79
Db 78 ALTSPAYNK-FDCGEMRSVQYTGYLEVRMKPAKNTGIVSFFTYGTPT--DQTPMDE 133

Qy 80 VDIEVLGNKPGSFQSNITITGKAGAKTSEKHHAVSPADQAFHTYTGLEMTPNYVRYRWYDGO 139
Db 134 IDIEFLGKDTTKVQFNYYTNGAG--NHEKIVDLGFDANANVHTYAFDMQPNISIKWYVDG 190

Qy 140 QEVRTGGVSNLTGTGLRFLNLSSEA-AWVGQFDESKLPLEQFINWYKVKYK 193
Db 191 Q-LKHTATNQTTPPK--IMMNLNMGTVDEWLGSYNGVN-PLYAHYDWRVRYK 241

RESULT 4
US-09-286-690-10
: Sequence 10, Application US/09286690
: Patent No. 6103511
: GENERAL INFORMATION:
: APPLICANT: Li, Xin-Liang
: APPLICANT: Ljungdahl, Lars G.
: APPLICANT: Chen, Huizhong
: TITLE OF INVENTION: Lichenase and Coding Sequences
: FILE REFERENCE: 55-96
: CURRENT APPLICATION NUMBER: US/09/286,690
: CURRENT FILING DATE: 1999-04-05
: EARLIER APPLICATION NUMBER: US 60/027,882
: EARLIER FILING DATE: 1999-10-04
: EARLIER APPLICATION NUMBER: PCT/US97/17811
: EARLIER FILING DATE: 1997-10-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 10
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Bacillus licheniformis
: US-09-286-690-10

Query Match 12.8%; Score 233.5; DB 3; Length 243;
Best Local Similarity 33.7%; Pred. No. 1.5e-13;
Matches 58; Conservative 26; Mismatches 77; Indels 11; Gaps 6;

Qy 23 TNVSAKDFSGALVLTLEEVYQGFPEARKMAASGVSMFLYONGSEIADGRPWVEVDI 82
Db 81 TSPSYNKFDCGEMRSVQYTGYLEVRMKPAKNTGIVSFFTYGTPT--DQTPMDEIDI 137

Qy 83 EYLGNKPGSFQSNITITGKAGAKTSEKHHAVSPADQAFHTYTGLEMTPNYVRYRWYDGOEV 142
Db 138 EFLGKDTTKVQFNYYTNGV--NHEKIVNLGFDANANVHTYAFDMQPNISIKWYVDG-L 193

Qy 143 RKTGGVSNLTGTGGLRFLNLSSEA-AWVGQFDESKLPLEQFINWYKVKYK 193
Db 194 KHTATTQIPQTPK--IMMNLNMGAGVDEWLGSYN-GVTPLSRSLHWRYRK 242

RESULT 5
US-09-286-690-9
: Sequence 9, Application US/09286690
: Patent No. 6103511
: GENERAL INFORMATION:
```



```
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 279
; TYPE: PRF
; ORGANISM: Clostridium thermocellum
US-09-286-690-9
```

```
Query Match          12.7%; Score 231.5; DB 3; Length 279;
Best Local Similarity 35.3%; Pred. No. 2,7e-13;
Matches 61; Conservative 23; Mismatches 70; Indels 19; Gaps 7;
```

```
QY 28 KDFSGAELYTLEEVQ-----YGFPEARKMAAASGTVSSMFLYONGSEIADGRPWVEVD 81
DB 82 REYGSYPRKSGEYRTKSEFGYEVRRMAKNVGISSFTYTGPS--DNNPDEID 138
QY 82 IEVLGNPSPFSQNIITGKAGAKTSEKHHAIVSPADQAFHTTGLMTPRYVWTVDGCE 141
DB 139 IEFLGDTTKRVQFNMWKNVGG--NEYLHNLGFDASODFHTGTFGRPDYIDFYVDGKK 195
QY 142 VRTEGGOVSNLTGTG-LRFNLMSSESA-AWVGOPDESKLPFLFOFINWVKVY 192
DB 196 VVR-----GTRNIPVTPGKIMNMLPBGIGVDEWLGRD-GRTPLOAEYEVKYY 243
```

```
RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 238
; TYPE: PRF
; ORGANISM: Bacillus polymyxa
US-09-286-690-7
```

```
Query Match          12.5%; Score 227.5; DB 3; Length 238;
Best Local Similarity 34.3%; Pred. No. 5e-13;
Matches 61; Conservative 26; Mismatches 74; Indels 17; Gaps 9;
```

```
QY 20 AITTVASADEFGAELYTLEEVQYGFPEARKMAAASGTVSSMFLYONGSEIADGRPWVE 79
DB 74 SLTSPANNK-FCCGGEYRSTNNYGYGLYEVSMKPAKNTGIVSSFTYTGPSH--GTOMDE 129
QY 80 VDEVLGNKPNPSPFSQNIITGKAGAKTSEKHHAIVSPADQAFHTTGLMTPRYVWTVDG 139
DB 130 IIEFLGDKTTKRVQRYVTVNGVGG---EKIINLGFDASTSFHTTAFFDWOPGIRKMYVDG 186
QY 140 QEVKTEGGVSNLTGTG-LRFNLMSSESA-AWVGOPDESKLPFLFOFINWVKVYKT 195
```

```
DB 187 -VLKHT---ATTNIPSTPGKIMNMLNMGTVDSMLGSIYNAN-PLVAFYDWV---KYT 236
```

```
RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 242
; TYPE: PRF
; ORGANISM: Clostridium thermocellum
US-09-286-690-11
```

```
Query Match          12.3%; Score 224.5; DB 3; Length 242;
Best Local Similarity 33.5%; Pred. No. 9,5e-13;
Matches 62; Conservative 25; Mismatches 71; Indels 27; Gaps 8;
```

```
QY 14 VAAAAAALTITNV-----ANDFGAELYTLEEVQ-----YGFPEARKMAAASGTV 59
DB 58 VSTVLAFTGIDISNKKMLTLDREYGSYPRKSGEYRTKSEFGYEVRRMAKNVGI 117
QY 60 SSMFLYONGSEIADGRPWVEIEVLGNPSPFSQNIITGKAGAKTSEKHHAIVSPADQ 119
DB 118 SSFTYTGPS--DNNPDEIDIEFLGKDTTKRVQFNMWKNVGG--NEYLHNLGFDASO 171
QY 120 AFHTTGLEWTPNVYRWTVDGQEVKRTGEGOVSNLTGTG-LRFNLMSSESA-AWVGOPDE 177
DB 172 DFHTYGFEMRPDYIDFYVDGKKVYR-----GTRNIPVTPGKIMNMLPBGIGVDEWLGRD- 226
QY 178 SKPL 182
DB 227 GRPL 231
```

```
RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
; GENERAL INFORMATION:
; APPLICANT: Borrijs, Rainer
; APPLICANT: Hofemeister, Jurgen
; APPLICANT: Thomsen, Karl Kristian
; APPLICANT: Olsen, Ole
; APPLICANT: Von Wettstein, Dietrich
; TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 1800 Diagonal Road, P.O. Box 299
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/103,998
: FILING DATE:
: CLASSIFICATION: 435
: PRIORITY APPLICATION NUMBER: US 07/773,652
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Bent, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30307/123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-9300
: TELEFAX: (703) 683-4109
: TELEX: 899149
: INFORMATION FOR SEQ. ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 237 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-103-998-4
```

```
Query Match          12.2%; Score 222.5; DB 1; Length 237;
Best Local Similarity 33.1%; Pred. No. 1.4e-12;
Matches 57; Conservative 25; Mismatches 79; Indels 11; Gaps 6;
```

```

OY 23 TNYSADESGAELTYLEEVGYKFEARMKMAASGTYSMFLYONGSEIADGRPMVEVDI 82
DB 75 TSSAYNKFDCAERSTNYGYGLYVSMKPAKNTGLVSSFFTYTG- --AHGTOMBEIDI 131
OY 83 EYLKKNPGSFOSNIITGKGAQKTSKHNHVASPADQAFHTYGLWTPTNYVRMTVDGQEV 142
DB 132 EFLGKDTTKVQFNYYNGAG--NHEKFDLGFDAANAYTFAFMDQPNISKYVDSQ-L 187
OY 143 RKTEGGVSNLUTGCIREFLMSSES-AMVGFDESKLPLFOFINMKVYK 193
DB 188 KHTATTOIPAAPGK--IMMNLMTGTVDMGLSGSYNGVN-PIYAHYDMRYRK 236
```

```

RESULT 9
US-09-286-690-2
: Sequence 2, Application US/09286690
: Patent No. 6103511
: GENERAL INFORMATION:
: APPLICANT: Li, Xin-Liang
: APPLICANT: Ljungdahl, Lars G.
: APPLICANT: Chen, Huizhong
: TITLE OF INVENTION: Lichenase and Coding Sequences
: FILE REFERENCE: 55-96
: CURRENT APPLICATION NUMBER: US/09/286,690
: CURRENT FILING DATE: 1999-04-05
: EARLIER APPLICATION NUMBER: US 60/027,882
: EARLIER FILING DATE: 1999-10-04
: EARLIER APPLICATION NUMBER: PCT/US97/17811
: EARLIER FILING DATE: 1997-10-03
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: Patentin Ver. 2.0
: SEQ. ID NO 2
: LENGTH: 245
: TYPE: PRT
: ORGANISM: Orplomyces sp. PC-2
: US-09-286-690-2
```

```
Query Match          10.4%; Score 190.5; DB 3; Length 245;
Best Local Similarity 31.6%; Pred. No. 1.2e-09;
Matches 54; Conservative 17; Mismatches 87; Indels 13; Gaps 6;
```

```

OY 21 LATTNVSADPSGAELTYLEEVGYKFEARMKMAASGTYSMFLYONGSEIADGRPMVEV 80
DB 79 LTIIDRGSGTGTGEYRTKNTKNGYGMVNMKPKLNKNGYVSSFFTYTGPS--DGTWDEI 135
OY 81 DIEVLGKNPGSFOSNIITGKGAQKTSKHNHVASPADQAFHTYGLWTPTNYVRMTVDGQ 140
```

```

DB 136 DIEFLGDTTKVQFNYYNGAGH---EHIHYLGFDASGFFHTYGFEMRNSITWYDGT 192
OY 141 EVKRTGGGVSNLGTGQ-LREFLMSSES-AMVGFDESKLPLFOFINMV 189
DB 193 AVYTA---YDNIPTPGKIMNMANNGIGVDDMLRPPN-GRINISAYIDWV 238
```

```

RESULT 10
US-08-824-707-2
: Sequence 2, Application US/08824707
: Patent No. 5919688
: GENERAL INFORMATION:
: APPLICANT: Ferrer, Pau
: APPLICANT: Diers, Ivan
: APPLICANT: Hedegard, Lisbeth
: APPLICANT: Halkier, Torben
: APPLICANT: Asenjo, Juan
: TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 59196880 No. 5919688disk of No. 5919688th America, Inc.
: STREET: 405 Lexington Avenue, Suite 6400
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25 (Epo)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/824,707
: FILING DATE: 14-April-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Gregg, Valera A.
: REGISTRATION NUMBER: 35,127
: REFERENCE/DOCKET NUMBER: 4290.204-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ. ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 306 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-824-707-2
```

```
Query Match          8.7%; Score 159.5; DB 2; Length 306;
Best Local Similarity 27.6%; Pred. No. 1.1e-06;
Matches 59; Conservative 29; Mismatches 89; Indels 37; Gaps 11;
```

```

OY 2 NIKTKAVKSALAVAAAAALTTVNSAKDFSGAELTYLEEV--QYKFEARMKMAASGTIV 59
DB 103 NYTTSRVNSALDQCGALVITALDQSDGTSYTSALTTQGNVQFQFGIETARIQIIPROGIM 162
OY 60 SSMFLYONGSEIADGRPMV--EVDI-EVLGNP---GSFOSNIITGKGAQKTSSEKHN 111
DB 163 SAFWNV--GANLPD-TPMPTSGEIDIMENVGNAPHEVHGTIVHGPVSGDNGIMGTVOHQ 219
OY 112 AVSPADQAFHTYGLWTPTNYVRMTVDGQEVKRTGEGVSNLGTGGLRFLNLMSSSAAW 171
DB 220 GMSFADD--FHTFGIDMTGCEITMLVDGQEHYHRTYADV-----ANQM 261
OY 172 VGOFDKSLPLFOFINMKVYKTYTPQGGGSGDF 205
DB 262 V--FDO---PFLILN-VAIGGOMFGNPDAITTF 289
```



```

Oy 56 SGTVSMHLXONGSGIADGRWV---EVDI-EVLGNP-----GSFQSIITGKAGNOKTS 107
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 148 OG-IPAPWMLGGS---PPGTPMPSSGIEDIMENNGFEPHRYHGVHGPXSGSGI--TG 2020
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 108 EKHNAPDAADQAFHTVGLSEMTPIYVRYMTYDGEVRKTEBGSQVSNLGTQGLRNLMSSE 1676
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 MYOHQNGSFDATPTEFTEFVADMRPELITWFDVGDOGFHRYTRASVG----- 246
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 168 SAAMWAGOFDESKLPLPFOFINWKKYVKTPCGGECS-----DFTLMDTNDNFTEDGS 219
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 247 ANAWV--FDQ---PEFLILN-VAVGGCMPIGPDDTTLDDPOMKADYVRYVINDSGS 296
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14  
US-08-712-072C-4  
; Sequence 4, Application US/08712072C

1 Patent NO. 5925541  
 2 GENERAL INFORMATION:  
 3 APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng  
 4 TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE  
 5 NUMBER OF SEQUENCES: 13  
 6 CORRESPONDENCE ADDRESS:  
 7 ADDRESSSEE: Amster, Rothstein & Ebenstein  
 8 STREET: 90 Park Avenue  
 9 CITY: New York  
 10 STATE: NY  
 11 COUNTRY: US  
 12 ZIP: 10016  
 13 COMPUTER READABLE FORM:  
 14 MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE  
 15 COMPUTER: IBM PC COMPATIBLE  
 16 OPERATING SYSTEM: MS-DOS  
 17 SOFTWARE: ASCII  
 18 CURRENT APPLICATION DATA:  
 19 APPLICATION NUMBER: US/08/712.072C  
 20 FILING DATE: 11-SEP-1996  
 21 CLASSIFICATION: 435  
 22 PRIOR APPLICATION DATA:  
 23 APPLICATION NUMBER:  
 24 FILING DATE:  
 25 ATTORNEY/AGENT INFORMATION:  
 26 NAME: Bogosian, Elizabeth A.  
 27 REGISTRATION NUMBER: 39,911  
 28 REFERENCE/DOCKET NUMBER: 63475/97  
 29 TELECOMMUNICATION INFORMATION:  
 30 TELEPHONE: (212) 697-5995  
 31 TELEFAX: (212) 286-0854 or 286-0082  
 32 TELEX: TWX 710-581-4766  
 33 INFORMATION FOR SEQ ID NO: 4:  
 34 SEQUENCE CHARACTERISTICS:  
 35 LENGTH: 276 amino acids  
 36 TYPE: amino acid  
 37 STRANDEDNESS:  
 38 TOPOLOGY: linear  
 39 MOLECULE TYPE: peptide  
 40 HYPOTHEICAL: NO  
 41 ANTI-SENSE: NO  
 42 FRAGMENT TYPE:  
 43 ORIGINAL SOURCE:  
 44 ORGANISM: gud, Rhodothermus marinus  
 45 US-08-712-072C-4

Query Match	7.38	Score	133.5	DB	2	Length	276
Best Local Similarity	23.78	Pred. No.	0.00023				
Matches	44	Conservative	34	Mismatches	85	Indels	23
						Gaps	8

[illegible]

```
Oy 18 DGGGKKEGGGVSLTC-----TGLGRNLMSSSAAWVG--PDESKPLPGFINW 188
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 DDSLRYRRPNRNLTPREADMRHMPDPFHLIMNIAYGANGGOQGYDPEAFPAQLVYDI 268
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 189 VKVKRY 194
      | : : : : :
Db 269 VVVVYW 274
```

RESULT 15  
US-08-712-072C-3

```

? Sequence 3, Application US/08712072C
? Patent No. 5925541
? GENERAL INFORMATION:
?   APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
?   TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
?   NUMBER OF SEQUENCES: 13
?   CORRESPONDENCE ADDRESS:
?     ADDRESSSEE: Amster, Rothstein & Edeinstein
?     STREET: 90 Park Avenue
?     CITY: New York
?     STATE: NY
?   COUNTRY: US
?   ZIP: 10016
? COMPUTER READABLE FORM:
?   MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
?   COMPUTER: IBM PC COMPATIBLE
?   OPERATING SYSTEM: MS-DOS
?   SOFTWARE: ASCII
?   CURRENT APPLICATION DATA:
?     APPLICATION NUMBER: US/08/712.072C
?     FILING DATE: 11-SEP-1996
?     CLASSIFICATION: 435
?     PRIOR APPLICATION DATA:
?       APPLICATION NUMBER:
?       FILING DATE:
?     ATTORNEY/AGENT INFORMATION:
?       NAME: Bogosian, Elizabeth A.
?       REGISTRATION NUMBER: 39,911
?       REFERENCE/DOCKET NUMBER: 63475/97
?     TELECOMMUNICATION INFORMATION:
?       TELEPHONE: (212) 697-5995
?       TELEFAX: (212) 286-0854 or 286-0082
?       TELEX: TWX 710-581-4766
?     INFORMATION FOR SEQ ID NO: 3:
?       SEQUENCE CHARACTERISTICS:
?         LENGTH: 321 amino acids
?         TYPE: amino acid
?         STRANDEDNESS:
?         TOPOLOGY: linear
?         MOLECULE TYPE: peptide
?         HYPOTHEetical: NO
?         ANTI-SENSE: NO
?         FRAGMENT TYPE:
?         ORIGINAL SOURCE:
?           ORGANISM: el3b, Bacillus circulans
?
? US-08-712-072C-3

```

Query Match	6.58;	Score 119;	DB 2;	Length 321;
Best Local Similarity	20.48;	Pred. NO. 0.006;		
Matches 43; Conservative	34;	Mismatches 84;	Indels 50;	Gaps 7.

```

0Y 20 ALTTVNSAKD-----FSGAEYLTLIE--VOYGFTEAMKMAAASGVYSSMFLYONSEI 71
Db 14 AMNDSKSPQDPNRYAOYSSGRIKNTDKJSLKLYGRVDFPAKJLPTGDGWPALMMLPKDSYV 183
0Y 72 ADGRPMVEVDI-EVLGNKPGSFOSNIIING---KAGAKTSEKHHAVSPADQAFHTYGLE 127
Db 184 GTMAASGELDVNEAGRLPCSVSGIITHGCGMPVNOSSGGDYHHPREGTFFANDHYHSYV 244
0Y 128 WTPNRYVRMTVDGQEVARKTEGGVSNLJTGQGLRPNLWSSESA-----165

```

Thu Jan 9 12:09:17 2003

us-09-654-652a-3.ra1

Page 7

Db 244 WEEDNIKKYVDGKFFKYVTNQ-----WYSTAAPNNPAPDEPFYIMNL 289

Qy 170 AMWGQFD-----ESKLPLOQFINWVKYKY 194  
| | | | |  
| | | | |  
290 AVGCGFDGRTPNASDIPATQVQVYVHYKF 320

Search completed: January 9, 2003, 12:06:35  
Job time : 15.7338 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:49:36 : Search time 9.75926 Seconds  
(without alignments)  
747.688 Million cell updates/sec

Title: US-09-654-652A-1  
Perfect score: 1333  
Sequence: 1 MVSADFGALTYLLEEVY.....TRKGSEFNGQVPRDEPAP 248

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1103	82.7	228	3	US-09-286-690-12
2	245.5	18.4	242	3	US-09-286-690-8
3	244	18.3	239	3	US-08-103-998-2
4	231.5	17.4	279	3	US-09-286-690-9
5	229.5	17.2	243	3	US-09-286-690-10
6	225	16.9	238	3	US-09-286-690-7
7	222.5	16.7	242	3	US-09-286-690-11
8	220.5	16.5	237	1	US-08-103-998-4
9	189	14.2	245	3	US-09-286-690-2
10	154.5	11.6	306	2	US-08-824-707-2
11	145	10.9	263	4	US-09-159-106-2
12	143	10.7	435	4	US-09-159-106-11
13	140.5	10.5	303	4	US-09-159-106-13
14	133.5	10.0	276	2	US-08-712-072C-4
15	118	8.9	321	2	US-08-712-072C-3
16	117.5	8.8	289	2	US-08-580-545B-4
17	117.5	8.8	289	4	US-09-262-653A-4
18	117.5	8.8	289	4	US-08-867-484A-2
19	117	8.8	620	4	US-09-269-731-6
20	116.5	8.7	654	1	US-08-392-828C-2
21	116.5	8.7	654	3	US-09-330-945-2
22	115	8.6	262	1	US-08-392-828C-37
23	115	8.6	262	3	US-09-330-945-37
24	113	8.5	285	2	US-08-712-072C-5
25	113	8.5	545	4	US-09-269-731-8
26	110.5	8.3	295	1	US-08-295-657-3
27	110	8.3	287	4	US-08-640-737-8

28	109.5	8.2	289	4	US-08-640-737-2	Sequence 2, Appl1
29	105	7.9	422	2	US-08-712-072C-2	Sequence 2, Appl1
30	102.5	7.7	395	2	US-08-404-531B-3	Sequence 3, Appl1
31	102.5	7.7	395	3	US-08-476-900A-3	Sequence 3, Appl1
32	102.5	7.7	395	3	US-08-488-546A-3	Sequence 3, Appl1
33	102.5	7.7	1580	4	US-08-726-320-1	Sequence 1, Appl1
34	102.5	7.7	1580	4	US-09-208-716-1	Sequence 1, Appl1
35	102.5	7.7	1581	4	US-08-726-320-3	Sequence 3, Appl1
36	102.5	7.7	1581	4	US-09-208-716-3	Sequence 3, Appl1
37	95.5	7.2	1581	2	US-08-404-531B-6	Sequence 6, Appl1
38	95.5	7.2	1581	3	US-08-476-900A-6	Sequence 6, Appl1
39	95.5	7.2	1581	3	US-08-488-546A-6	Sequence 6, Appl1
40	95.5	7.2	1581	4	US-08-726-320-4	Sequence 4, Appl1
41	95.5	7.2	1581	4	US-09-208-716-4	Sequence 4, Appl1
42	95.5	7.2	1582	2	US-08-404-531B-9	Sequence 9, Appl1
43	95.5	7.2	1582	3	US-08-476-900A-9	Sequence 9, Appl1
44	95.5	7.2	1582	3	US-08-488-546A-9	Sequence 9, Appl1
45	95.5	7.2	1582	4	US-08-726-320-5	Sequence 5, Appl1

# ALIGNMENTS

```

RESULT 1
US-09-286-690-12
; Sequence 12, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; EARLIER FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-286-690-12

Query Match      82.7%  Score 1103:  DB 3:  Length 228:
Best Local Similarity 100.0%  Pred. No. 3e-106:
Matches 204;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

OY  2  VSADDFGALTYLLEEVYQKFEARKMAASGVSMFLYONGSEIADGRPWEVDIEV 61
    |||||||
Db  25  VSADDFGALTYLLEEVYQKFEARKMAASGVSMFLYONGSEIADGRPWEVDIEV 84
    |||||||

OY  62  LCKNPSPFOSNITTTGKAGAKTSEKHHAVSPRADOAFHTYTGLEMTNRYRWYTDGGEVRK 121
    |||||||
Db  85  LCKNPSPFOSNITTTGKAGAKTSEKHHAVSPRADOAFHTYTGLEMTNRYRWYTDGGEVRK 144
    |||||||

OY  122  TEGGQVSNLTGTOGLFFNLMSSESAAWGOFPDESKLPLFOFIWVYKYVTPGCGGGSDD 181
    |||||||
Db  145  TEGGQVSNLTGTOGLFFNLMSSESAAWGOFPDESKLPLFOFIWVYKYVTPGCGGGSDD 204
    |||||||

OY  182  FTLDWTDNFDTEGSRWKGKDWTF 205
    |||||||
Db  205  FTLDWTDNFDTEGSRWKGKDWTF 228
    |||||||

RESULT 2
US-09-286-690-8
; Sequence 8, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:

```

APPLICANT: Li, Xin-Liang  
APPLICANT: Ljungdahl, Lars G.  
APPLICANT: Chen, Huizhong  
TITLE OF INVENTION: Lichenase and Coding Sequences  
FILE REFERENCE: 55-96  
CURRENT APPLICATION NUMBER: US/09/286,690  
CURRENT FILING DATE: 1999-04-05  
EARLIER APPLICATION NUMBER: US 60/027,882  
EARLIER FILING DATE: 1999-10-04  
EARLIER APPLICATION NUMBER: PCT/US97/17811  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO: 8  
LENGTH: 242  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
US-09-286-690-8

Query Match 18.4%; Score 245.5; DB 3; Length 242;  
Best Local Similarity 35.8%; Pred. No. 1.5e-17;  
Matches 59; Conservative 24; Mismatches 71; Indels 11; Gaps 6;

QY 7 FSGAELTYLEEVQYKFEARKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLKNP 66  
DB 87 FDGGRSVQYTGGLYEVRKPAKNTGIVSSFFTYTGPT---DGTPEIDIEFLGKDT 143  
QY 67 GSFQSNITIGKAGAKTSEKHNHVASPADQAFHTYGLWTPNYPVMTVDGQEVKRT 126  
DB 144 TKVQFNVTNGAG---NHEKIDLGFDANAHHTVAFDMQPNISKYVDSQ-LKHTATNQ 199  
QY 127 VSNLTGTGRLFNLMSSESA-AWVGQFDESKLPLFQFINVKYK 170  
DB 200 IPTTPGK--IMNMLNGTGVDEWLGSYNGVN-PLVAHYDWRVTK 241

RESULT 3  
US-08-103-998-2  
Sequence 2, Application US/08103998  
Patent No. 5470725  
GENERAL INFORMATION:  
APPLICANT: Borliss, Rainer  
APPLICANT: Hofemeister, Jürgen  
APPLICANT: Thomsen, Karl Kristian  
APPLICANT: Olsen, Ole  
APPLICANT: Von Wettstein, Dietrich  
TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 1800 Diagonal Road, P.O. Box 299  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/103,998  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/773,652  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30307/123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 239 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-103-998-2

Query Match 18.3%; Score 244; DB 1; Length 239;  
Best Local Similarity 35.5%; Pred. No. 2.1e-17;  
Matches 61; Conservative 26; Mismatches 69; Indels 16; Gaps 8;

QY 3 SAKDFSGAELTYLEEVQYKFEARKMAASGVSSMFLYONGSEIADGRPWVEVDIEVL 62  
DB 80 SYNKFDGGRSVQYTGGLYEVRKPAKNTGIVSSFFTYTGPT---GTPNDEIDIEFL 136  
QY 63 GKNPGSFQSNITIGKAGAKTSEKHNHVASPADQAFHTYGLWTPNYPVMTVDGQEVKRT 122  
DB 137 GKDTTKVQFNVTNGVGH---EKVISLGFDAKSGFHTVAFDMQPGYIKMYVDG-VLKHT 192  
QY 123 EGGQVSNLTGTGQ-LRFNLMSSESA-AWVGQFDESKLPLFQFINVKYK 172  
DB 193 ---ATANIPSTPGKIMNMLNGTGVDDWLGSYNGAN-PLVAEDWV---KYT 237

RESULT 4  
US-09-286-690-9  
Sequence 9, Application US/09286690  
Patent No. 6103511  
GENERAL INFORMATION:  
APPLICANT: Li, Xin-Liang  
APPLICANT: Ljungdahl, Lars G.  
APPLICANT: Chen, Huizhong  
TITLE OF INVENTION: Lichenase and Coding Sequences  
FILE REFERENCE: 55-96  
CURRENT APPLICATION NUMBER: US/09/286,690  
CURRENT FILING DATE: 1999-04-05  
EARLIER APPLICATION NUMBER: US 60/027,882  
EARLIER FILING DATE: 1999-10-04  
EARLIER APPLICATION NUMBER: PCT/US97/17811  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO: 9  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Clostridium thermocellum  
US-09-286-690-9

Query Match 17.4%; Score 231.5; DB 3; Length 279;  
Best Local Similarity 35.3%; Pred. No. 5.2e-16;  
Matches 61; Conservative 23; Mismatches 70; Indels 19; Gaps 7;

QY 5 KDFSGAELTYLEEVQ-----YKFEARKMAASGVSSMFLYONGSEIADGRPWVEVD 58  
DB 82 REGGSYPRKSGEYRTRKSPFGYGYVRKMAKKNQIVSSFFTYGPT---DNNPWEID 138  
QY 59 IEVLGKNPGSFQSNITIGKAGAKTSEKHNHVASPADQAFHTYGLWTPNYPVMTVDGQ 118  
DB 139 IEFGLKDTTKVQFNMYKKNQVGC---NEYLNLGFPASQDFHTYGFEMRPDYIDFYVDCK 195  
QY 119 VRKTEGQVSNLTGTGQ-LRFNLMSSESA-AWVGQFDESKLPLFQFINVKY 169  
DB 196 VYR---GTRNIPVTPGKIMNMLNGTGVDDWLGRYD-GRPLQAEYEVYKY 243

RESULT 5  
US-09-286-690-10  
Sequence 10, Application US/09286690  
Patent No. 6103511  
GENERAL INFORMATION:



```

; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRF
; ORGANISM: Bacillus licheniformis
US-09-286-690-10

Query Match
Best Local Similarity 17.2%; Score 229.5; DB 3; Length 243;
Matches 57; Conservative 25; Mismatches 76; Indels 11; Gaps 6;

OY 3 SAKDFGAEIYTLAEVOYGFKFAKMAAASGVSSMFLYQNGSEIADGRPWVEVDIEVL 62
DB 84 SYKFFDCGERSVQYTGGLYEYVMKPAKNVGVSSFFTYTGP--DGTMPDEIDIEFL 140
OY 63 GKRPGSFOSIITGKGAQKTSKHNHVASPADQAFHTYGLFMTPTVYKRTVDSQGEYKRT 122
DB 141 GKTTVVOFYNYNGVG---NHKIVNLGFDANSHYTVAFDMQPSIKMYVDGQ--LKHT 196
OY 123 EGGQVSNLGTGGLRFNLMSSESA-AMVGQFDESKLPLFOFIMVVKYK 170
DB 197 ATTQIQTQTEK--IMMNLNAGVDEMLGSYN-GVPLRSLSLHWVRYTK 242

RESULT 6
US-09-286-690-7
; Sequence 7, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 238
; TYPE: PRF
; ORGANISM: Bacillus polymyxa
US-09-286-690-7

Query Match
Best Local Similarity 16.9%; Score 225; DB 3; Length 238;
Matches 58; Conservative 23; Mismatches 71; Indels 16; Gaps 8;

OY 7 FSGAEIYTLAEVOYGFKFAKMAAASGVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
DB 83 FDCGEYRSTNNYVGLYEYVMKPAKNVGVSSFFTYTGP--DGTMPDEIDIEFLGKDT 139
OY 67 GSPOSIIITGKGAQKTSKHNHVASPADQAFHTYGLFMTPTVYKRTVDSQGEYKRTGCG 126
DB 140 TKYQFNYNYNGVG--EKLINLGFDASTSFHTYAFDMQPSIKMYVDG--VLKHT--A 192
OY 127 VSNLTGTG--LRFNLMSSESA-AMVGQFDESKLPLFOFIMVVKYK 172
```

```

DB 193 TTNIPSTPGKIMMNLNMGCTVDMSLGSYNGAN-PIVAEYMW---KYT 236

RESULT 7
US-09-286-690-11
; Sequence 11, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 242
; TYPE: PRF
; ORGANISM: Clostridium thermocellum
US-09-286-690-11

Query Match
Best Local Similarity 16.7%; Score 222.5; DB 3; Length 242;
Matches 58; Conservative 22; Mismatches 64; Indels 19; Gaps 7;

OY 5 KDFSGAEIYTLAEVQ-----YKFAKMAAASGVSSMFLYQNGSEIADGRPWVEVD 58
DB 80 REYGGSYPRKSGEYRTKSPFGYVEYVRMAKKNVGVSSFFTYTGPS--DNNPDEID 136
OY 59 IEVLGKNPFSOSIITGKGAQKTSKHNHVASPADQAFHTYGLFMTPTVYKRTVDSQGE 118
DB 137 IEFLGKDTTKVQFNMYKNYGVG---NEYLNILGFDASQDPHTTYGFEKRPDYIDFYVDGK 193
OY 119 VRTEGGQVSNLGTGGLRFNLMSSESA-AMVGQFDESKLPL 159
DB 194 VYR---GTRNIPVTPGKIMMNLMPGIGVDEWLGRD-GRTPL 231

RESULT 8
US-08-103-998-4
; Sequence 4, Application US/08103998
; Patent No. 5470725
; GENERAL INFORMATION:
; APPLICANT: Bottis, Rainer
; APPLICANT: Hofmeister, Jurgen
; APPLICANT: Thomsen, Karl Kristian
; APPLICANT: Olsen, Ole
; APPLICANT: Von Wettstein, Dietrich
; TITLE OF INVENTION: A Thermostable (1,3-1,4)-beta-Glucanase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: FOLEY & LARDNER
; STREET: 1800 Diagonal Road, P.O. Box 299
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/103,998
; FILING DATE:
; CLASSIFICATION: 435
```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/773,652  
FILING DATE: 16-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30307/123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-103-998-4

Query Match 16.5%; Score 220.5; DB 1; Length 237;  
Best Local Similarity 33.9%; Pred. No. 5, 6e-15;  
Matches 56; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

QY 7 FSGAEYLTLEEVQYGFKFAKMAAAGTVSSMFLYONGSEIADGRPWEVDIEVLGNP 66  
DB 82 FDCAEYRSTNIYGLYEVSMKPAKNTGIVSFFTYGP---AHGTQWDEIDIEFLGKDT 138  
QY 67 GSFQSNITTKAGAKTSEKHNHVASPRADAFTHTGLEMTPNVRVTVDDQEVKRTGCG 126  
DB 139 TKVQFNYYNAG--NHEKFAADGFDAAVAYHTYAFDWQPSIKWVDQ-LKHTATTQ 194  
QY 127 VSNLTGTQGLRFNLMSSSA-AWVGOFDESKLPLFOFINNVKVK 170  
DB 195 IIPAPCK--IMNMLNMGTVGVDMLGSGNGVN-PIYAHYDMKRRK 236

RESULT 9  
US-09-286-690-2  
Sequence 2, Application US/09286690  
Patent No. 6103511  
GENERAL INFORMATION:  
APPLICANT: Li, Xin-Liang  
APPLICANT: Ljungdahl, Lars G.  
APPLICANT: Chen, Huizhong  
TITLE OF INVENTION: Lichenase and Coding Sequences  
FILE REFERENCE: 55-96  
CURRENT APPLICATION NUMBER: US/09/286,690  
EARLIER FILING DATE: 1999-04-05  
EARLIER APPLICATION NUMBER: US 60/027,882  
EARLIER FILING DATE: 1999-10-04  
EARLIER APPLICATION NUMBER: PCT/US97/17811  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 245  
TYPE: PRT  
ORGANISM: *Orpiniomyces* sp. PC-2  
US-09-286-690-2

Query Match 14.2%; Score 189; DB 3; Length 245;  
Best Local Similarity 32.9%; Pred. No. 1, 1e-11;  
Matches 56; Conservative 16; Mismatches 76; Indels 22; Gaps 8;

QY 5 KDSGAEYLTLEEVQ-----YKFEARKMAAAGTVSSMFLYONGSEIADGRPWEVD 58  
DB 83 RDSGSG---YTCGEYRTKKNYVGMFQVNMKPIKNPGVSSFFTYGPS--DGTKDEID 136  
QY 59 IEVLGNPFSFOSNITTKAGAKTSEKHNHVASPRADAFTHTGLEMTPNVRVTVDDQEV 118  
DB 137 IEFLGIDITTKVQFNYYNAGGHH---EHITVLGFDASQGHHTGFWARNSITWYVDGA 193  
QY 119 VRKTEGGOVSNLTGTG-LRFNLMSSSA-AWVGOFDESKLPLFOFINNV 166

DB 194 VYTA-----YDNIDPTGKIMNMAMNGIGVDDWDLRPFN-GRTNISAYVDWV 238

RESULT 10  
US-08-824-707-2  
Sequence 2, Application US/08824707  
Patent No. 5919688  
GENERAL INFORMATION:  
APPLICANT: Ferrer, Pau  
APPLICANT: Diers, Ivan  
APPLICANT: Hedegaard, Lisbeth  
APPLICANT: Halkier, Torben  
APPLICANT: Asenjo, Juan  
APPLICANT: Savva, Demitris  
TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 59196880 No. 5919688disk of No. 5919688th America, Inc.  
STREET: 405 Lexington Avenue, Suite 6400  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/824,707  
FILING DATE: 14-April-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gregg, Valecia A.  
REGISTRATION NUMBER: 35,127  
REFERENCE/DOCKET NUMBER: 4290,204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 306 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-824-707-2

Query Match 11.6%; Score 154.5; DB 2; Length 306;  
Best Local Similarity 28.5%; Pred. No. 5, 5e-08;  
Matches 53; Conservative 28; Mismatches 68; Indels 37; Gaps 11;

QY 7 FSGAEYLTLEEV--OYKFEARKMAAAGTVSSMFLYONGSEIADGRPV---EVDI-E 60  
DB 131 YTSARLTGQNVQPOFGRIEARIQIPRCQIGWSAFWV--GANLPD-TWPTSGEIDME 187  
QY 61 VLGNP-----GSFQSNITTKAGAKTSEKHNHVASPRADAFTHTGLEMTPNVRVTVG 116  
DB 188 NGSMAHEVHGVHGPYSGDNGIMCTQHPGQWSPFAD--FTFEIDMTPEGITWLVG 245  
QY 117 QEVKTEGGOVSNLTGTGLRFNLMSSSAWVGOFDESKLPLFOFINNVKVKYKPGOG 176  
DB 246 QEVHRTVTDVG-----ANQWV--FDQ---PFLLLN-VALIGGWRGPN 283

QY 177 EGGSDF 182  
DB 284 DATTPF 289  
RESULT 11  
US-09-159-106-2  
Sequence 2, Application US/09159106  
Patent No. 6284509

```

: GENERAL INFORMATION:
: APPLICANT: Ferrer, Pau
: APPLICANT: Diers, Ivan
: APPLICANT: Halkier, Torben
: APPLICANT: Hedegaard, Lisbeth
: TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
: TITLE OF INVENTION: Activity
: FILE REFERENCE: 4693..204-US
: CURRENT APPLICATION NUMBER: US/09/159,106
: CURRENT FILING DATE: 1998-09-23
: EARLIER APPLICATION NUMBER: 0427/96
: EARLIER FILING DATE: 1996-12-04
: EARLIER APPLICATION NUMBER: 0885/96
: EARLIER FILING DATE: 1996-08-23
: EARLIER APPLICATION NUMBER: PCT/DK97/00160
: EARLIER FILING DATE: 1997-04-14
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 263
: TYPE: PRT
: ORGANISM: Oerskovia xanthineolytica
: OS-09-159-106-2

```

[illegible]

```

RESULT 12
US-09-159-106-11
: Sequence 11, Application US/09159106
: Patent No. 6284509
: GENERAL INFORMATION:
: APPLICANT: Ferrer, Pau
: APPLICANT: Diers, Ivan
: APPLICANT: Halkier, Torben
: TITLE OF INVENTION: An Enzyme With -1,3-Glucanase
: TITLE OF INVENTION: Activity
: FILE REFERENCE: 4693.204-US
: CURRENT APPLICATION NUMBER: US/09/159,106
: CURRENT FILING DATE: 1998-09-23
: EARLIER APPLICATION NUMBER: 0427/96
: EARLIER FILING DATE: 1996-12-04
: EARLIER APPLICATION NUMBER: 0885/96
: EARLIER FILING DATE: 1996-08-23
: EARLIER APPLICATION NUMBER: PCT/DK97/00160
: EARLIER FILING DATE: 1997-04-14
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 435
: TYPE: PRT
: ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

```

Query Match	10.7%	Score 143:	DB 4:	Length 435:
Best Local Similarity	24.6%	Pred. No. 1.4e-06:		
Matches 62:	Conservative 29:	Mismatches 95:	Indels 66:	Gaps 14:

  

QY	19	QYGRKEAFRMKMAAGSGVSSMFLYONGSEIADGRPV---	EVDI-EVLGKNP---	GSFQ 70
		:     : : :		
Db	134	QYGRLEAIIQIPIRGCG-IPWAFMYLGG--	PPGTPWPSGGELIDMENNGFEDHRKHYVH	190
		:     : : :		
QY	71	SNITGKGAQKTSSEKHNAVSPAADQAFHYTGLEMTYVRYWTVYDGOEYRKTGGQVSNL		130
		:     : : :		
Db	191	GRGVSGSGSI--TGKXGHPQCGSFADTPTHTFAVDMKPEIITMFVYDGGQGFHRYTRASVG-		246
		:     : : :		
QY	131	TGTGTCGRFLNLMSSF-----	SAAMVGQFD-ESKLPLEQLTNWKKVYKYPFGGCEG	178
		:     : : :		
Db	247	-----ANAWFCDPFFLLINVAVGCGMPGPDGTTQLDPOOMKYDYRVYDNGSGSSSP		299
		:     : : :		
QY	179	GSDFT-----	LDWTFNFD-----	TFGDS-----
		:     : : :		
Db	300	GNPGLPTGTGAVFAANGMCVDVPMADPTDGNPNQIVITCSGNAAQWTTRGS--	DGT-V	355
		:     : : :		
QY	211	DLTDENIYSRDG 222		
		:     : : :		
Db	356	RALGKCLDYRDG 367		

```

RESULT 13
US-09-159-106-13
: Sequence 13, Application US/09159106
: Patent No. 6284509
: GENERAL INFORMATION:
: APPLICANT: Ferrer, Pau
: APPLICANT: Diers, Ivan
: APPLICANT: Halkier, Torben
: APPLICANT: Heidegaard, Lisbeth
: TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
: TITLE OF INVENTION: Activity
: FILE REFERENCE: 4693..204-US
: CURRENT APPLICATION NUMBER: US/09/159,106
: CURRENT FILING DATE: 1998-09-23
: EARLIER APPLICATION NUMBER: 0427/96
: EARLIER FILING DATE: 1996-12-04
: EARLIER APPLICATION NUMBER: 0885/96
: EARLIER FILING DATE: 1996-08-23
: EARLIER APPLICATION NUMBER: PCT/DK97/00160
: EARLIER FILING DATE: 1997-04-14
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: SeqSeq for Windows Version 3.0
: SEQ ID NO 13
: LENGTH: 303
: TYPE: PRT
: ORGANISM: Oerskovia xanthoneolytica
US-09-159-106-13

```

Query Match	10.5%	Score 140.5;	DB 4;	Length 303;
Best Local Similarity	27.4%	Pred. No. 1.5e-06;		
Matches	52;	Conservative	25;	Mismatches 74; Indels 39; Gaps 11;
QY	19	QYKTEARKKMAAASGVTSMEFLYONGSEIDADGRWV---	EVDI-EVLGKNP---	GSFQ 70
		:     :     :		
Db	134	QYRIEARLQIPRQSG-IMPAFWMLGGS--PRGRTWPSGSEIDIMENVGFEEDHRYHGTIV		190
QY	71	SNITTKGKGAQCKTSKKNHNAVSFADQADHTYTGLENTPRVYKRTVYGVGSEVRYKTEGGVSNL		130
		:     :     :		
Db	191	GRGVSGSGGI--TGNYQIPQGWSEADDTHTFAVDKPRPELITWFDVGQGFVHRTASVQ--		246
QY	131	TGTQGLRFNINLMSSESAANVGQFDESKRLPFOFINNVKYYKTPGQGEESG---	DTFLDM	186
		:     :     :		
Db	247	-----ANAMV--FDQ--PFLLIN-VAVGGQWPGVRDTGTLQRPQMKVDY		286
QY	187	TDNEDPTFDGS	196	
		:   :		
Db	287	VRVYDNGSGS	296	

```
RESULT 14
US-08-712-072C-4
; Sequence 4, Application US/08712072C
; Patent No. 5925541
; GENERAL INFORMATION:
; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rochstein & Ebenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,072C
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A.
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: gub, Rhodothermus marinus
; US-08-712-072C-4

Query Match
Best Local Similarity 23.7%; Pred. No. 76-06; Length 276;
Matches 44; Conservative 34; Mismatches 85; Indels 23; Gaps 8;

OY 5 KDFSGAELTYLEEVQ--YGFKFEARKMAAASGVSSMFLYONGSEIADGRPV---EVDI 59
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 93 REYTSARLYTRGKASWTYGFREIRARLPSCGTWPAIMWLPDRQYGSAY-WPDNGEIDI 151
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 60 -EVLGNP-----GSFOSNITTKAGAKQKSEKHNHAYSPAAQAFHTYGLEWTPNYRYRWY 114
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 152 MEHVAGNPDPVHGTHTKAYNHLGTRGGSIR--VPTARTDFHYAYALEMTEPEELRMFV 208
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 115 DCGEVRKTEGGOVSNLTG-----TGGLRFNLMSSESAWGO--FDESKLPLPQFINW 165
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 209 DDSLTYRFPNERLITDEADNRHMPFDDPFLIMINIVAGGAMGGOGGVDPAPPAQLVVDY 268
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 166 VKVYKY 171
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 269 VRVYRW 274
```

```
RESULT 15
US-08-712-072C-3
```

```
; Sequence 3, Application US/08712072C
; Patent No. 5925541
; GENERAL INFORMATION:
; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rochstein & Ebenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,072C
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A.
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: e13b, Bacillus circulans
; US-08-712-072C-3

Query Match
Best Local Similarity 20.5%; Pred. No. 0.00035; Length 321;
Matches 40; Conservative 31; Mismatches 80; Indels 44; Gaps 6;

OY 7 FSGAELTYLEF--VOYGFKFEARKMAAASGVSSMFLYONGSEIADGRVWEYDI-EVLG 63
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 140 YSSGKINTKDKLSLKRGVDFRAKLPTGDGWPMLPRKDSYGTWASGEIDVDEARG 199
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 64 KNGSFOSNITTS---KAKAQKSEKHNHAYSPAAQAFHTYGLEWTPNYRYRWTVDSQEV 120
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 200 RLEGSVSGTILHEGGQRPVNOSSGDDYHFPREGQTFANDYIVYSVWMEEDIKRYVDGKFFY 259
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 121 KTEGGGVSNLTGGLRFNLMSSESA-----ANVGQFD-----ESK 156
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 260 KVTNQ-----WYSTAAPNPNPAFPDEPFYLLIMNLAVGNGFDDGGRTPMASD 305
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 157 LPLFQFINWVYKY 171
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 306 IPATMQVDYVRYKYF 320
```

```
Search completed: January 9, 2003, 12:06:32
Job time : 10.7593 secs
```



Qy	61	SMFLYONSELAIDORPMWVEVIEVLGNKPGSFOSNIITGKGAOKTSEKHHAVSPADQA	1200
Db	61	SMFLYONSELAIDGRPMWVEVIEVLGNKPGSFOSNIITGKGAOKTSEKHHAVSPADQA	1200
Qy	121	FHTYGLEMTPMVVMYVDGOEVRKTEGGVSNLTGTGLRENLWSSSEAMAVGOFDESKL	180
Db	121	FHTYGLEMTPMVVMYVDGOEVRKTEGGVSNLTGTGLRENLWSSSEAMAVGOFDESKL	180
Qy	181	PLFOPINNVKYYKXTTGGGEGGSGFTLDMTMDNPTFGGSRMGKCDMTFFDCNRYDLTDKNI	240
Db	181	PLFOPINNVKYYKXTTGGGEGGSGFTLDMTMDNPTFGGSRMGKCDMTFFDCNRYDLTDKNI	240
Qy	241	YSRDGMILLALTRKGOESFNGQVPRDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV	300
Db	241	YSRDGMILLALTRKGOESFNGQVPRDEPAPQSSSSAPASSSSVPASSSSVPASSSSAFV	300
Qy	301	PPSSSSATNALTHGMRTTPAVAKKEHRNLVNAKGAUVNNGHKRYRVVNEH	349
Db	301	PPSSSSATNALTHGMRTTPAVAKKEHRNLVNAKGAUVNNGHKRYRVVNEH	349

```

: RESULT 2
: US-09-791-30441
: Sequence 30441, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomimix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 30441
: LENGTH: 349
: TYPE: PRT
: ORGANISM: Fibrobacter succinogenes
: US-09-791-30441

```

Query Match	100.0%	Score 1824	DB 21	Length 349	
Best Local Similarity	100.0%	Pred. No. 1,3e-151			
Matches 349	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	MNIKTTAVKSAALAVAAAAAALTNTNYSAKDFSGAEHLYTLEEYQYGRKFPEAKMKAASGTVS	60		
Db	1	MNIKTTAVKSAALAVAAAAAALTNTNYSADDFSGAEHLYTLEEYQYGRKFPEAKMKAASGTVS	60		
QY	61	SMFLYQNSGSEIADGRPWVEVDIEVLGCKNPDSQSNITGKACAQKSTSEKHAIVSPAADOA	120		
Db	61	SMFLYQNSGSEIADGRPWVEVDIEVLGCKNPDSQSNITGKACAQKSTSEKHAIVSPAADOA	120		
QY	121	FHTYGLEMTPNVVRVTYVGQEVKRTKEGQVSNLTGQGRFLMLKSSSAWAGQDESEKL	180		
Db	121	FHTYGLEMTPNVVRVTYVGQEVKRTKEGQVSNLTGQGRFLMLKSSSAWAGQDESEKL	180		
QY	181	PLFQFINNVKVKYKYPGQGESESDTLDMTDNDFDFDSRMRCKGMDTFPGNRVDLTDKNI	240		
Db	181	PLFQFINNVKVKYKYPGQGESESDTLDMTDNDFDFDSRMRCKGMDTFPGNRVDLTDKNI	240		
QY	241	YSRDGMILLALTRKGQESFNGQVPRDDEPAPOSSSSAPASSSSVPAASSSSVPAASSSAFV	300		
Db	241	YSRDGMILLALTRKGQESFNGQVPRDDEPAPOSSSSAPASSSSVPAASSSSVPAASSSAFV	300		
QY	301	PPSSSSATNALHGMRTTPAVAKEHENLVNAGAKKPNPNCBKRYRNFEN	349		
Db	301	PPSSSSATNALHGMRTTPAVAKEHENLVNAGAKKPNPNCBKRYRNFEN	349		

```

1  GENERAL INFORMATION:
2  APPLICANT: SHYUR, LIE-FEN
3  APPLICANT: CHEN, JUI-LIN
4  APPLICANT: YANG, NING-SUN
5  TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLAST SUCCINOGENES 1, 3-1,
6  TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMAITC ACTIVITY AND
7  TITLE OF INVENTION: THERMO-TOLERANCE
8  FILE REFERENCE: 4910-8
9  CURRENT APPLICATION NUMBER: US/09/654,652A
10 CURRENT FILING DATE: 2000-09-05
11 NUMBER OF SEQ ID NOS: 11
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO 2
14 LENGTH: 267
15 TYPE: PRT
16 ORGANISM: Artificial Sequence
17 FEATURE:
18 OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
19 OTHER INFORMATION: with enhanced activity and thermal stability
20 US-09-652A-2

```

	Query Match	73.5%	Score 1340;	DB 20;	Length 267;
	Best Local Similarity	99.6%	Pred. No. 4	1e-109;	
	Matches 250;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	25	VSAKDFSCAELTYLLEEVYQKFEAPRMKMAASGVYSSMFLYQNSEIADGRPWEEVDIEV	84		
Db	2	VSAKDFSCAELTYLLEEVYQKFEAPRMKMAASGVYSSMFLYQNSEIADGRPWEEVDIEV	61		
OY	85	LGNKNGSFQSNITITKACAOKTSEKHNAVSPAADAFITTYGEMTPPNVVRMTYQGEYRK	144		
Db	62	LGNKNGSFQSNITITKACAOKTSEKHNAVSPAADAFITTYGEMTPPNVVRMTYQGEYRK	121		
OY	145	TEGGCVSNULTGTQGLRFLNMSSESAAWVGQFDESKLPLPQFINWKKYKKTTPGQEGGSD	204		
Db	122	TEGGCVSNULTGTQGLRFLNMSSESAAWVGQFDESKLPLPQFINWKKYKKTTPGQEGGSD	181		
OY	205	FTLDWTDNDFDFDGSGRMGKDMTFPFGNVDLTDKNIYSRDMLILALTRKQESFNGVP	264		
Db	182	FTLDWTDNDFDFDGSGRMGKDMTFPFGNVDLTDKNIYSRDMLILALTRKQESFNGVP	241		
OY	265	RDDEPAPQSSS 275			
Db	242	RDDEPAPQSSS 252			

RESULT 4  
 US-09-654-652A-1  
 : Sequence 1, Application US/09654652A  
 : GENERAL INFORMATION:  
 : APPLICANT: SHYR, LIE-FEN  
 : APPLICANT: CHEN, JUI-LIN  
 : APPLICANT: YANG, NING-SUN  
 : TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLAST SUCCINOGENES 1, 3-1,  
 : TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND  
 : TITLE OF INVENTION: THERMO-TOLERANCE  
 : FILE REFERENCE: 4910-8  
 : CURRENT APPLICATION NUMBER: US/09/654,652A  
 : CURRENT FILING DATE: 2000-09-05  
 : NUMBER OF SEQ ID NOS: 11  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO: 1  
 : LENGTH: 248  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme  
 : OTHER INFORMATION: with enhanced activity and thermal stability  
 : US-09-654-652A-1

RESULT 3  
US-09-654-652A-2  
; Sequence 2, Application US/09654652A

Query Match	72.8%;	Score 1328;	DB 20;	Length 248;
Best Local Similarity	100.0%;	Pred. No. 4.2e-108;		
Matches 247;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	25	VSADFCSAELTYLLEBYOYGFEEARMMAAASGVSSMPLYONSSETADBRPWYEV	84
Db	2	VSADFCSAELTYLLEBYOYGFEEARMMAAASGVSSMPLYONSSETADGPRWYEV	61
Qy	85	LGNKNGSFOSNIITGKGAOKTSEKHHAVSPAADOAEHTGLETPNYVRYMTVDGQEVK	144
Db	62	LGNKNGSFOSNIITGKGAOKTSEKHHAVSPAADOAEHTGLETPNYVRYMTVDGQEVK	121
Qy	145	TEGGOVSNLTGQGLRFLNLSSESAANVGOFPDESKLPLFOFINWVKYKXTTPOGEGGSD	204
Db	122	TEGGOVSNLTGQGLRFLNLSSESAANVGOFPDESKLPLFOFINWVKYKXTTPOGEGGSD	181
Qy	205	FTLDPMIDNFDFFDSRMKGKGMTEFDGNRVYLTDKNINYSRQGMILATLRKGOSEFNQVP	264
Db	182	FTLDPMIDNFDFFDSRMKGKGMTEFDGNRVYLTDKNINYSRQGMILATLRKGOSEFNQVP	241
Qy	265	RDDPEAP 271	
Db	242	RDDPEAP 248	

```

RESULT 5
US-09-791-537-145019
: Sequence 145019, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 145019
: LENGTH: 214
: TYPE: PRT
: ORGANISM: pdb 1CPM
US-09-791-537-145019

```

Query Match	14.9%;	Score 272.5;	DB 21;	Length 214;
Best Local Similarity	33.8%;	Pred. No. 3.5e-15;		
Matches 77; Conservative	33;	Mismatches 93;	Indels 25;	Gaps 12

[illegible]

```

RESULT 6
US-09-791-537--106873
: Sequence 106873, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210

```

```

CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 106873
LENGTH: 259
TYPE: PRT
ORGANISM: Brevibacillus brevis
US-09-791-537-106873

```

Query Match	14.6%;	Score 266.5;	DB 21;	Length 259;
Best Local Similarity	-36.0%;	Pred. NO. 1.6e-14;		
Matches 63;	Conservative 24;	Mismatches 75;	Indels 13;	Gaps 6

[illegible]

```

RESULT 7
US-09-791-537-32335
; Sequence 32335, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32335
;
; LENGTH: 252
;
; TYPE: PRT
;
; ORGANISM: Brevibacillus brevis
US-09-791-537-32335

```

Query Match	14.6%;	Score 265.5;	DB 21;	Length 252;
Best Local Similarity	-35.8%;	Pred. NO. 1.8e-14;		
Matches 62;	Conservative 24;	Mismatches 74;	Indels 13;	Gaps 6

Qy 26 SAKPDSGAELETIEVOYGRKEARKMMAAGCTYSSMEFLVONCEISIDGRNVEIDIEVL 85  
 Db 90 SARNTKAGELTNTPIYHGLEFVSMKPAKVEGTYSSFFLY-TGSDMDGDGPMDIEDIEFL 148  
 Qy 86 GKNGSGPDSNIITGKAGAKQKTSKHHAVSPAADAFTHTGLEMTPNRYKRTVTVDGQEVKRT 145  
 Db 149 GKDTTRIOFNFTFTGVGG--NEFTYIDGDFPDSFPTIAPAEKEDSDITMTVNGEAAHTA 205  
 Qy 146 EGGGVSNLTGT-QGLRNLWSSESA-AMVGOEFSCKLPLEQFINMKVRYKRT 196  
 Db 206 ---TENIPQPRIMNMLMGVGVDAQTGTFDDCDPNVSYSDMV---RYTP 251

```

RESULT 8
US-09-463-862A-1
: Sequence 1, Application US/09463862A
: GENERAL INFORMATION:
: APPLICANT: Hillen, Wolfgang
: APPLICANT: Maurer, Karl-Heinz
: TITLE OF INVENTION: New Beta-Glucanase from a Bacillus
: FILE REFERENCE: H2031PCT/US
: CURRENT APPLICATION NUMBER: US/09/463,862A

```





Db 80 TSPSYNFECDGSGNSVOTYGYGLYEVMKPAKNGIYSSFFYTGP --- DGPWDEIDI 136

OY 83 EYLKNGSPQSNIITCKAGOKTSEKHNHVASPAADQAFYTYGLEMPTVRNRYTVOGEV 142  
| | | | |  
Db 137 EFLKOTITKYQVFNNTYGAG --- NHEIYVDLGRDAANAAYITPAQWQPSNIKRYVDSQ-L 152

OY 143 KRTGGOVSULTGIGLRFNLMSESA -AVNGOEDSKLPLFOFINWKKYK 193

Db 193 KHTATNDIPTTPGK - IMMLNMGSTGVDELGLSYNGVN -PLVAYHYDVRATK 241

RESULT 12  
US-09-791-537-103347

Db 77 TSPYNNFDDGGENRSVQTYGCLYEVMKAKRNTGIYSSFFTYTGPT---GTFPMDEID 133

QY 83 EVLGKNGSFOSNITITGKAGAOKTSEKHAAVSPADAQFHTYGLWETPNYVRMTYDQEV 142

Db 134 EFLCKDPTTKVOFNYYTNGVSGH--EKVISLGFASGKFTHTYAFDMQPGYIKRWYDQ-VL 189

QY 143 RKTEGGGVSNLTGQS--LRNIMSSSEA-AMWQGFDESKLPLOQFIMMYVYVYT 195

Db 190 KHT--ATAINPTSPGKIMNMIMNGICVDMDLGSYGAN-PLTAEDYMW---KYT 237

```

RESULT 14
US-09-791-537-73450
; Sequence 73450, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73450
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-09-791-537-73450

```

Query Match	13.6%	Score 248;	DB 21;	Length 242;
Best Local Similarity	36.0%	Pred. No. 6, 1e-13;		
Matches	63;	Conservative	25;	Mismatches 75;
			Indels	12;
			Gaps	7

  

QY	20	ALTITVNSAKDFSCAELTYLLEEVYQKFEARKMAAASCTVSMFLYONGSEIADGRPWE	79
DB	78	ALTSAPVANK-FDCGERSVQTYGTGYEVRMKPAKNTGIVSFFLYTGP	133
QY	80	VDIEVLCKNPGSQSNITGKAKAQKTSKHNHASPAAADQAFHTYGLFMPPIYKRWTVDG	139
DB	134	IDIEFLGKDPYTKQFNVYNTGAG---NHKIKYIDLGFDAANAHYHFAEPMQPNISKRWYDG	190
QY	140	QEVKRTGEGGVSLTGTGGLRFINIWSBEA-AWVGQPDSEKLPLEQFINNVKYYK	193
DB	191	Q-LKHTATNDIPTGK--IMNLMNGTGVDLGLSTYGVN-PLTAHYDWRYTK	241

```

RESULT 15
US-09-791-537-79889
: Sequence 79889, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomimix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 79889
: LENGTH: 242
: TYPE: PRT
: ORGANISM: Bacillus subtilis
: US-09-791-537-79889

```

Query Match	13.68;	Score 248;	DB 21;	Length 242;
Best Local Similarity	36.08;	Pred. No. 6.1e-13;		
Matches 63;	Conservative 25;	Mismatches 75;	Indels 12;	Gaps 7

```

QY 20 ALTTNVSARKDFSGAELTYLLEVOYKFEARKMMAASGTVSSMFLYQNGSEIADGRPWVE 79
    ||: | | | : : | | | : | | | : | | | : | | |
Db 78 ALTSPAYNK-FDCGENRSVQTYGYGLYVRMKPAKNTGISSFFTYTGP--DGTWDE 133
    : | | | : | | | : | | | : | | | : | | |
QY 80 VDIEVLGKNPGSFQSNITIGKAGAKTSEKHHA VSPADQAFHTYGLEWTPNIVRNTVDG 139
    : | | | : | | | : | | | : | | | : | | |
Db 134 IDIEFLGKDPYKQFNYYTNGAG--NHEKIVDLGFDANANAYHTYAFDQPNISIKWYVDG 190
    : | | | : | | | : | | | : | | | : | | |
QY 140 OEVRKTEGGOVSNLTGGLRFLNLSSES-AWVGOFDESKLPLFQFINWVKYK 193
    | : : | | : | | : | | : | | : | | : | | : |
Db 191 Q-LKHTATNQIPTTPGK--IMNLMNGTGVDEMLGSYNGVN-PLIAHYDWRITK 241
    | : : | | : | | : | | : | | : | | : | | : |

```

Search completed: January 9, 2003, 12:13:29  
 Job time : 165.402 secs







A:Residues:1-334 <SCH>  
A:Cross-references: EMBL:X6335; NID:g40697; PIDN:CA44959.1; PID:g40698  
C:Genetics:  
A:Gene: *llcB*  
C:Superfamily: licheninase *llcB*; Clostridium cellulase repeat homology  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-334/Product: licheninase #status predicted <Mat>  
F:273-296/Domain: Clostridium cellulase repeat homology <CCR1>  
F:308-331/Domain: Clostridium cellulase repeat homology <CCR2>

Query Match	13.08;	Score 237;	DB 1;	Length 334;
Best Local Similarity	30.98;	Pred. No. 1.4e-09;		
Matches 80; Conservative 27; Mismatches 94; Indels 58; Gaps 13				

```

QY 28 KDSGSAELTYLLEEVQ-----YGFEEARMKMAAASGVSSMFLYONSELSADCRPAPEVVD 81
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 82 REYGGSTPYRKSGEELRTYSFEGYGYEYRMKAANKVGVLSVFSEFTYTGDS---DNNPDEID 138
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 82 IEVLGKNQGSFOSNIITGKAGAAOKTSEKHNHAVSPAADAQHTHTGLEMTPNYVWMTVDGGE 141
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 139 IEFLGKDTTKTVQFMWYKNGVGS---NEYLNHLGPDASQDHPHTGYGFEKRPIDYIDFYDGKK 195
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 142 VRKREGGVSNULTTQGL-LRPNLMSSSSA-AWVGQFPEESKLPLFQFLNNYKVY----- 192
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 196 VYR---GTNRNIPVTPPKIMNNLWPGIGVDWELGRVD-GRTPLOAEYEVKKYPNGVPOD 250
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 193 -----KYTP-----GQGE-GGSDF-----LDMFDNFDPFDSRMKG 224
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 251 NPTPTPIATPSTPIPNPLPLKGDVNGDGHVNSSDYSLFKRYLLRVLDIFRVYGQS---VA 307
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 225 DMTEDGKRVLDLTKNIYSR 243
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 308 DVNRDG-KIDSTDLTLMKR 325
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 9  
S15388  
licheninase (EC 3.2.1.73) - *Bacillus licheniformis*  
N:Alternate names: endo-beta-1,3-1,4-D-glucanase; lichenase  
C:Species: *Bacillus licheniformis*  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S15388  
R:Robbers, J.; Perez-Pons, J.A.; Querol, E.  
Eur. J. Biochem. 197, 337-343, 1991  
A:Title: Molecular cloning, expression and nucleotide sequence of the endo-beta-1,3-1,4  
#:Reference number: S15388; MUID:91224124; PMID:2026156

Query Match	12.8%;	Score 233.5;	DB 1;	Length 243;
Best Local Similarity	33.7%;	Pred. NO. 1.7e-09;		
Matches 58;	Conservative 26;	Mismatches 77;	Indels 11;	Gaps 6;

OY 23 TNSAKDPSCAELYTELVNOVGKFEARPMKAASGVYSKFLYONGSEIADGRPWEVDI 82  
 : : : : : : : : : : : : : : : : : :  
 Db 81 TSPYNKRFDCGENSEVTVGYGLYEVMKPAKNVGIASFVSFTYTGP7---DGPMDIEDI 137  
 : : : : : : : : : : : : : : : : : :  
 OY 83 EVLCKNGSGFOSNIITTKACAQKTSEKHNAVSPAADOAFHTTGLEMTPNVVRKVTVDGEV 142  
 : : : : : : : : : : : : : : : : : :  
 Db 138 EFLCKDTTKVQFNYYTTNGVC---NHEIKVALIGFDAANSYHTVAFDQMOPNSIKWYVDGQ-L 193  
 : : : : : : : : : : : : : : : : : :  
 OY 143 RKTGGCVSNLTGTGRLRNLMSSESA-AANVGPFDESKLP.LPFOFIWMVKYK 193  
 : : : : : : : : : : : : : : : : : :  
 Db 194 KHTPTTQIPOTPRK--IMMLDMNGAGVDMLGSYN-GVPLSLSLIHWVRTK 242

## RESULT 10

S64507  
probable membrane protein YGR189c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein G7553

A:Residues: 1-507 <ARR>  
A:Cross-references: EMBL:Z79974, NID:g1323335, PID:e243566, GSPDB:GN0000  
A:Experimental source: strain 5288C

A;Gene: SCD:CRH1; MIPS:YGR189C  
A;Cross-references: SGD:S0003421  
A;Map position: 7R  
C;Keywords: transmembrane protein  
F;6-22/Domain: transmembrane #sta

Query Match	12.6%;	Score 230;	DB 2;	Length 507;
Best Local Similarity	26.9%;	Pred. No. 7.5e-09;		
Matches	87;	Conservative 57;	Mismatches 134;	Indels 46;
				Gaps 13;

```

0Y 6 TAVKSAALVAAAAAALTNNKSKDSSGAEIVLTLEVOYCKEAREBKMAAAGTSVSMFLY 65
Db 70 TDLKHAIGEIKYSGDGLSMTL -AKRIDNPSLSNFIYIMIGKLEVIILKANGTGIYVSFLQ 120
0Y 66 ONGSEIADGRPWVEVDIEVLCKNPGSFOSNIITYGAKAQKTSKHHAAVSPADQAFTHTYG 123
Db 129 SDDL- -----EIDIEWGGDNTOQFOSNFESKGDITTYRGEHGHVDPFDK -FHNVT 178
0Y 126 LEWTFNYVRWTFVDGGOEVRTEGCGVSNLTGTGG-----LRFNLMS- ---ESA- ---W 178
Db 180 LDMANDKTTWYLDGSSVR-----VLSNTSSGCIYQSPMYLTMGIMAGCDPDNNACTIEW 238
0Y 172 V-GQFDESKLPLEOFINWVKYKYKYPGGGEGSDFTLDMTDFTFDGSRWKCGDMTFDG 238
Db 234 AGGETNYNDAPFTWIEKVIYLDVYSTGKKRYTGDDSGSM -ESIENDGSGISYGRYDOAE- 298
0Y 231 NRVDLTIDKNIYSRDMILLALTRKG -QESFNGCVPRDDEPAPOSSASAPASS---SSVPA 280
Db 232 -----DRAVLANGSISSSSTSSSTVSSASASTYSSSVSSSTVSSASASTYSSSVSTVS 348
0Y 287 SSSSVPASSSSSAEVPSSSSSATNA 310
Db 346 SSSSVSSSSSST- -PSSSTATSS 366

```

RESULT 11  
S11927  
licheninase (EC 3.2.1.73) precursor [validated] - *Bacillus macerans*  
N:Alternate names: endo-beta-1,3-1,4-gluconase; lichenase  
C:Species: *Bacillus macerans*  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2000  
C:Accession: S11927  
R:Roriss, R.; Buettner, K.; Maentsaelae, P.  
Mol. Gen. genet. 222, 278-283, 1990  
A:Title: Structure of the beta-1,3-1,4-gluconase gene of *Bacillus macerans*: homologue  
A:Reference number: S11927; MUID:91109712; PMID:2274030  
A:Accession: S11927  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <BOR>  
A:Cross-references: EMBL:X55959; NID:g296715; PIDN:CA59426.1; PID:g296716  
R:Hahn, M.; Heinemann, U  
submitted to the Brookhaven Protein Data Bank, December 1994  
A:Reference number: A67074; PID:IMAC  
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 26-37, 'P', 39-2  
A:Note: recombinant form with residues 26-41 from *Bacillus amyloliquefaciens* express  
R:Keller, T.; Simon, O.; Roriss, R.; Heinemann, U.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5287-5291, 1993

A:Title: Molecular and active-site structure of a Bacillus 1,3-1,4-beta-glucanase.  
A:Reference number: A47562; MUID:93281743; PMID:8099449  
A:Contents: annotation, x-ray crystallography, 2.3 angstroms  
A:Note: recombinant form with residues 26-41 from Bacillus amyloliquefaciens expressed in  
C:Function:  
C:Description: catalyzes the hydrolysis of 1,4-beta-D-glucosidic bonds in lichenin and  
C:Superfamily: licheninase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
E:1-25/Domain: signal sequence #status predicted <PRO>  
E:26-237/Product: licheninase #status experimental <MA>  
E:55-84/Disulfide bonds: #status experimental  
E:120/Active site: Glu #status predicted

Query Match 12.1%; Score 221; DB 2; Length 246;  
 Best Local Similarity 33.7%; Pred. No. 1.3e-08;  
 Matches 58; Conservative 19; Mismatches 73; Indels 22; Gaps 8;

OY 30 FSGAEIYLTLEBYQYCKEAKRMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLGKNP 89  
 DB 91 YAGGETRSNNRRGYGLYRSMKPAKHIGVDSFSTGPPS--DNNPWDEIDIEFLGKDT 147  
 OY 90 GSFSQNIITGKAGCAOKTSEKH--HAVSPADQAFTYGLWTPNYVRMTVDGQEVRAKTE 146  
 DB 148 TEVQEFNYIYNGV-----KHELYLKLGFDAKSGFHTYGYIMEQNTIAMLVDSKEVYRA- 200  
 OY 147 GGOVSNLTGTG-LRFNLWSSESA-AWGOFDSEKLPFLQFINWVKVYKTP 196  
 DB 201 --TSNIPTHPGKVMNMLPGLGVDSMLGAYD-GYTPVKAYYVMA--MYNP 245

## RESULT 15

AB3055  
 endo-1,3-1,4-beta-glucanase exok [Imported] - Agrobacterium tumefaciens (strain C58, DuF  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
 C:Accession: AB3055  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, I.  
 et al.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Pery, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB3577; PMID:11743193  
 A:Accession: AB3055  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-263 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AAL44856.1; PID:g17742502; GSPPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: exok  
 A:Map position: linear chromosome  
 C:Superfamily: licheninase

Query Match 11.3%; Score 206; DB 2; Length 263;  
 Best Local Similarity 35.1%; Pred. No. 1.7e-07;  
 Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

OY 28 KDSGAEIYLTLEBYQYCKEAKRMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLGK 87  
 DB 89 RNPACGEIQTGRYGTYEARKKATGSLNSAFPTYIGPT--DKKPHDEIDIEFLGK 145  
 OY 88 NPGSFQSN-IITGKAGCAOKTSEKHNAVSPADQAFTYGLWTPNYVRMTVDGQEVRAKTE 146  
 DB 146 NTGKYNQVQNLNVIANKG---NEKLIVVEGADAGFNDYAFVNEPQRLRYVNGKLVH--- 198  
 OY 147 GGOVSNLT---GTGGLRPNLWSSESA-AWGOF 175  
 DB 199 --EVTDETKIPONAKRIFPSLWCTDPLKDMGKF 230

Search completed: January 9, 2003, 12:19:01  
 Job time : 18.5613 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:06:42 ; Search time 8.8657 Seconds

(without alignments)  
1628.889 Million cell updates/sec

Title: US-09-654-652a-3  
Perfect score: 1824  
Sequence: 1 MNIKRTAVKASALAAVAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1824	100.0	349	1 GUB_FIBSU	P17989 fibrobacter
2	266.5	14.6	259	1 GUB_BACBR	P37073 bacillus br
3	248	13.6	242	1 GUB_BACSU	P04957 bacillus su
4	242.5	13.3	239	1 GUB_BACAM	P07980 bacillus am
5	242.5	13.3	802	1 XYND_RUMFL	O53317 rumnococcu
6	237	13.0	334	1 GUB_CLOTM	P29716 clostridium
7	233.5	12.8	243	1 GUB_BACLI	P27051 bacillus li
8	230	12.6	507	1 YG46_YEAST	P53301 saccharomyc
9	228	12.5	237	1 GUB_PAPMA	P23304 paenibacill
10	227.5	12.5	238	1 GUB_PAPPO	P45797 paenibacill
11	197	10.8	269	1 EXOK_RHIME	P33693 rhizobium m
12	189.5	10.4	347	1 UTR2_YEAST	P32823 saccharomyc
13	157.5	8.6	269	1 MERS_ARATH	P24806 arabidopsis
14	133.5	7.3	286	1 GUD_RHOMR	P45798 rhodothermu
15	128	7.0	283	1 BRU1_SOYBN	P35694 glycine max
16	119	6.5	465	1 EXSH_RHIME	O33680 rhizobium m
17	117	6.4	397	1 CGKA_ALTCA	P43478 altheromonas
18	115	6.3	682	1 E13B_BACCI	P23303 bacillus ci
19	108.5	5.9	989	1 SERRA_PLARF	P13823 plasmidum
20	107.5	5.9	719	1 YBPA_STRPN	O04707 streptococc
21	107	5.9	962	1 YODO_YEAST	P10476 saccharomyc
22	104	5.7	962	1 GUNA_PSEFL	O08193 pseudomonas
23	103.5	5.7	1580	1 ACC8_HUMAN	O09428 homo sapien
24	102.5	5.6	879	1 GUNI_CLOTM	O02934 clostridium
25	102	5.6	465	1 EGIC_RHIME	O02392 rhizobium m
26	102	5.6	650	1 SC02_NEUCR	O01277 neurospora
27	101	5.5	691	1 VIV1_MAIZE	P26307 zea mays (m
28	100.5	5.5	1036	1 NIT2_NEUCR	P19212 neurospora
29	100	5.5	748	1 GUNC_PSEFL	O27403 pseudomonas
30	99	5.4	344	1 GUN4_TREPR	O14405 trichoderma
31	99	5.4	525	1 Y487_TREPR	O83500 treponema p
32	99	5.4	560	1 YDEN_ESCHL	P77318 escherichia
33	98.5	5.4	658	1 GUN3_FIBSU	P14250 fibrobacter

34	98.5	5.4	1655	1	OMPB_RICCN	O9Kk3 r outer mem
35	98	5.4	1845	1	Z236_HUMAN	O9u136 homo sapien
36	97.5	5.3	1385	1	C5A8_BACUD	O45760 bacillus th
37	97.5	5.3	1581	1	ACC8_CRICR	O09427 cricetus cr
38	97.5	5.3	1581	1	ACC8_RAT	O09429 ratius norv
39	97	5.3	1770	1	PMPC_CHLTR	O84419 chlamydia t
40	96.5	5.3	526	1	CAP_YEAST	P17555 saccharomyc
41	96	5.3	371	1	LIG2_PHACH	P49012 phenetochae
42	95.5	5.2	872	1	IP3L_RAT	P42335 ratius norv
43	95.5	5.2	1289	1	C5A8_BACUD	O45753 bacillus th
44	95	5.2	400	1	GUN5_BACAG	O85465 bacillus ag
45	94.5	5.2	451	1	MYBH_DICDI	P34127 dictyostell

## ALIGNMENTS

```

RESULT 1
GUB_FIBSU
ID GUB_FIBSU STANFARD: PRT: 349 AA.
AC P17989;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Mixed linkage beta-
DE glucanase) (Lichenase).
OS Fibrobacter succinogenes (Bacteroides succinogenes).
OC Bacteria: Fibrobacter/Acidobacteria group: Fibrobacter group:
OC Fibrobacter.
OX NCBI_TaxID=833;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC STRAIN=Isolate S85;
RX MEDLINE=90299807; PubMed=2193918;
RA Teather R.M., Erfle J.D.;
RT "DNA sequence of a fibrobacter succinogenes mixed-linkage
RT beta-glucanase (1,3-1,4-beta-D-glucan 4-glucanohydrolase) gene."
RL J. Bacteriol. 172:3837-3841(1990).
CC -I- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -I- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: M33676; AAA24896.1; -.
DR PIR: A44507; A44507.
DR HSSP: P23904; IAIK.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydroxylase; Glycosidase; Signal; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 349
FT ACT_SITE 79 79
FT ACT_SITE 83 83
FT DOMAIN 271 307
FT REPEAT 271 277
FT REPEAT 278 284
FT REPEAT 285 291
FT REPEAT 292 298
FT REPEAT 301 307
SQ
SEQUENCE 349 AA. 37737 MW; 16DC4F5BDFC578A CRC64;
Query Match 100.0%; Score 1824; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MNIKTTAVKSAALVAAAAALTTNVSADKDFSGAELTYLEVOYKFEARKMAAAGTYS 60
DB 1 MNIKTTAVKSAALVAAAAALTTNVSADKDFSGAELTYLEVOYKFEARKMAAAGTYS 60
OY 61 SMELYONGSEIADGRPMVEVDIEVLCKNPGSFOGSIITGKAGAKTSEKHAAVSPAADA 120
DB 61 SMELYONGSEIADGRPMVEVDIEVLCKNPGSFOGSIITGKAGAKTSEKHAAVSPAADA 120
OY 121 FHTYGLEMPNRYVMTVDQEVKRTGEGOVSNLTGQGLRFNLMSESAAMVQGFDESKL 180
DB 121 FHTYGLEMPNRYVMTVDQEVKRTGEGOVSNLTGQGLRFNLMSESAAMVQGFDESKL 180
OY 181 PLFOFINMVKVYKRTPGGEGSDFTLDWTDNFDGSRWCKGDMTDFDGNRDLTDKNI 240
DB 181 PLFOFINMVKVYKRTPGGEGSDFTLDWTDNFDGSRWCKGDMTDFDGNRDLTDKNI 240
OY 241 YSRDGMILIALTRKGOSEFNGOVPRDDEPAPOSSSAPASSSSVPASSSSAV 300
DB 241 YSRDGMILIALTRKGOSEFNGOVPRDDEPAPOSSSAPASSSSVPASSSSAV 300
OY 301 PPSSSATNAIHGRTTPAVAKHRNLVNAKGAKNVNGCHRRVNFH 349
DB 301 PPSSSATNAIHGRTTPAVAKHRNLVNAKGAKNVNGCHRRVNFH 349

RESULT 2
GUB_BACBR STANDARD: PRT: 259 AA.
ID GUB_BACBR STANDARD: PRT: 259 AA.
AC P37073;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
GN (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
OS BGLHB.
OC Bacillus brevis.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OX NCBI_TaxId=1393;
RN [1]
RP MEDLINE=93159752; PubMed=7763386;
RA Louw M.E., Reid S.J., Watson T.G.;
RT "Characterization, cloning and sequencing of a thermostable
RT endo-(1,3-1,4) beta-glucanase-encoding gene from an alkalophilic
RT Bacillus brevis."
RL Appl. Microbiol. Biotechnol. 38:507-513(1993).
CC -1- FUNCTION: HYDROLASES B-GLUCANS CONTAINING MIXED BETA-1,3 AND
CC BETA-1,4 LINKAGES. THIS ENZYME IS THERMOSTABLE; ITS OPTIMAL
CC TEMPERATURE IS NEAR 65-70 DEGREES CELSIUS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M84339; AAA22265.1; -
CC DR HSSP: P23904; IADK.
CC DR InterPro: IPR000757; Glyco_hydro_16.
CC DR Pfam: PF00722; Glyco_hydro_16; 1.
CC DR PRINTS: PR00737; GLHYDRLASE16.
CC DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC KW Hydrolase; Glycosidase; Signal.
CC FT SIGNAL 1 31 POTENTIAL.
CC FT CHAIN 32 259 BETA-GLUCANASE.

```

```

FT ACT_SITE 142 142 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 259 AA: 29960 MW: A63C09F281F5D13 C8C64.

Query Match 14.6%; Score 266.5; DB 1; Length 259;
Best Local Similarity 36.0%; Pred. No. 2.9e-12;
Matches 63; Conservative 24; Mismatches 75; Indels 13; Gaps 6;

OY 26 SAKDFSGAELTYLEVOYKFEARKMAAAGTYSMELYONGSEIADGRPMVEVDIEVL 85
DB 90 SARNTAGELRTNDFRHYLFEYSKPAKVEGTSSFTY-IGENDMDGDPDELDIEFL 148
OY 86 GRNPGSFOGSIITGKAGAKTSEKHAAVSPAADAFTHTYGLEMPNRYVMTVDQEVKRT 145
DB 149 GKDTFRIGFNFTNGVG---NEFYDLDGFDASESFNTYAFEMRDSITWYNGCAVHTA 205
OY 146 EGGOVSNLTGT-QGLRFNLMSESA-AMVQGFDESKPLFOFINMVKVYKRTPGQ 198
DB 206 ---TENIPQTPQKIMMNLMPGVGDWGTGVPDGNTPVSYSDV---RYPPLQ 253

RESULT 3
GUB_BACSU STANDARD: PRT: 242 AA.
ID GUB_BACSU STANDARD: PRT: 242 AA.
AC P04957;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
GN (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
OS BGLS OR BGL OR LIC5 OR N15B.
OC Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C120;
RX MEDLINE=84272222; PubMed=6087283;
RA Murphy N., McConnell D.J., Cantwell B.A.;
RT "The DNA sequence of the gene and genetic control sites for the
RT excreted B. subtilis enzyme beta-glucanase."
RL Nucleic Acids Res. 12:5355-5367(1984).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=HL-25;
RC Tezuka H., Yunki T., Yabuuchi S.;
RT "Construction of a beta-glucanase hyperproducing Bacillus subtilis
RT using the cloned beta-glucanase gene and a multi-copy plasmid."
RL Agric. Biol. Chem. 53:2335-2339(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8696509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT covering the lic and cel loci, and creation of a 177 kb contig
RL Microbiology 142:3113-3123(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219081; PubMed=7704256;
RA Wolf M., Gezzi A., Simon O., Boriss R.;
RT "Genes encoding xylan and beta-glucan hydrolysing enzymes in Bacillus
RT subtilis: characterization, mapping and construction of strains
RL Microbiology 141:281-290(1995).
RN [5]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=BR151;
RX MEDLINE=96178961; PubMed=8606172;
RA Schnetz K., Stuelke J., Gertz S., Krueger S., Kriegl M., Hecker M.,

```

DR InterPro; IPR000757; Glyco\_hydro.16.  
 DR Pfam; PF00404; Dockerin\_1; 2.  
 DR Pfam; PF00722; Glyco\_hydro.16; 1.  
 DR PRINTS; PR00737; GLHYDRLASE16.  
 DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_2.  
 DR PROSITE; PS00448; CLOS\_CELLULOSE\_RPT; 2.  
 DR Hydrolase; Glycosidase; Signal; Repeat.  
 KW Hydrolase; Glycosidase; Signal; Repeat.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 334 BETA-GLUCANASE.  
 FT ACT\_SITE 136 136 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 140 140 PROTON DONOR (BY SIMILARITY).  
 FT DOMAIN 252 269 PRO/THR-RICH (LINKER).  
 FT DOMAIN 273 331 2 x 24 AA APPROXIMATE REPEATS.  
 FT REPEAT 273 331 1.  
 FT REPEAT 308 331 2.  
 FT CONFLICT 304 334 OSVADVNRDGRDSTDLMLKRYLRAIPSL ->  
 FT PDDGGRDRVYDSGK (IN REF. 2 AND 3).  
 SQ SEQUENCE 334 AA; 37897 MW; 0837564E9726F281 CRC64;

Query Match 13.0%; Score 237; DB 1; Length 334;  
 Best Local Similarity 30.9%; Pred. No. 5,1e-10;  
 Matches 80; Conservative 27; Mismatches 94; Indels 58; Gaps 13;

OY 28 KDFSGAEIYTLLEVO-----YKFEARKMAAAGTSSMFLYONGSEIADGRPWVEVD 81  
 DB 82 REYGSYPRKSGEYRFRKSPFGYGYEVRMKAKNNGIVSSFFTYTPS---DNNPWEID 138  
 OY 82 IEVLKRNPSFOSNIITGAGAKTSEKHAVSPADQAFHYTGLEWTNRYKWTYDGOE 141  
 DB 139 IEFLKRDITKYVFNMYKNGVGG---NEYLHNGFASODFHYGFEWRDIDFYVDGKK 195  
 OY 142 VKTEGGQVSNLTGTGOG-LRFNLMSESA-AWVGOFDESKLPFGFINNWKY----- 192  
 DB 196 VYR-----GRRNIPYTPGKIMMLNPGICVDENLGRID-GRTEPLQAEYEVKYYPNKVPD 250  
 OY 193 -----KYTP-----GQGE-GGSDFP-----LDWTDNDFDTGDSRFGKG 224  
 DB 251 NPTPTIAPSTPTPNPLPLKGDVNGDGHVNSDSSLFRKYLRLVDRPVGDS---VA 307  
 OY 225 DWTFPGNRYDLTDKNYSR 243  
 DB 308 DVNRDG-RIDSTDLMLTKR 325

## RESULT 7

CUB\_BACLI STANDARD: PRT; 243 AA.  
 ID GUB\_BACLI P27051;  
 AC 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)  
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).  
 GN Bgl.  
 OS Bacillus licheniformis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OC NCBI\_TaxID=1402;  
 RX NCBI\_TaxID=1402;  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=91224124; PubMed=2026156;  
 RA Lloberas J., Perez-Pons J.A., Querol E.;  
 RA "Molecular cloning, expression and nucleotide sequence of the  
 RA endo-beta-1,3-1,4-D-glucanase gene from Bacillus licheniformis.  
 RA Predictive structural analyses of the encoded polypeptide.";  
 RL Eur. J. Biochem. 197;337-343(1991).  
 RL [2]  
 RP REVISIONS.  
 RA Querol E.;  
 RL Submitted (JUL-1991) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP MUTAGENESIS.  
 RP MEDLINE=92362869; PubMed=1354172;

RA Planas A., Juncosa M., Lloberas J., Querol E.;  
 RT "Essential catalytic role of Glu134 in endo-beta-1,3-1,4-D-glucan 4-  
 RT glucanohydrolase from B. licheniformis as determined by site-directed  
 RT mutagenesis.";  
 RL FEBS Lett. 308:141-145(1992).  
 RN [4]  
 RN MUTAGENESIS.  
 RX MEDLINE=94237863; PubMed=8182059;  
 RA Juncosa M., Pons J., Dot T., Querol E., Planas A.;  
 RT "Identification of active site carboxylic residues in Bacillus  
 RT licheniformis 1,3-1,4-beta-D-glucan 4-glucanohydrolase by  
 RT site-directed mutagenesis.";  
 RL J. Biol. Chem. 269:14530-14535(1994).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RX MEDLINE=96063718; PubMed=7589539;  
 RA Hahn M., Pons J., Planas A., Querol E., Heinemann U.;  
 RT "Crystal structure of Bacillus licheniformis 1,3-1,4-beta-D-glucan 4-  
 RT glucanohydrolase at 1.8-A resolution.";  
 RL FEBS Lett. 374:221-224(1995).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages  
 CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
 CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE  
 CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X57279; CAA40547.1; -  
 DR PIR: S15388; S15388.  
 DR PDB: 1GBG; 07-DEC-95.  
 DR InterPro: IPR000757; Glyco\_hydro.16.  
 DR Pfam: PF00722; Glyco\_hydro.16; 1.  
 DR PRINTS: PR00737; GLHYDRLASE16.  
 DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR Hydrolase; Glycosidase; Signal; 3D-structure.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 243 BETA-GLUCANASE.  
 FT ACT\_SITE 134 134 NUCLEOPHILE.  
 FT ACT\_SITE 138 138 PROTON DONOR.  
 FT DISULFID 61 90  
 FT MUTAGEN 51 51 D->N: 30% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 89 89 D->N: 85% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 92 92 E->Q: 3% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 105 105 E->Q: 50% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 133 133 D->N: 15% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 134 134 E->Q: 0.2% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 136 136 D->N: 0.5% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 138 138 E->Q: COMPLETE LOSS OF ACTIVITY.  
 FT MUTAGEN 143 143 D->N: 65% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 160 160 E->Q: 15% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 168 168 D->N: 60% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 179 179 D->N: 80% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 190 190 D->N: 70% OF WILD-TYPE ACTIVITY.  
 FT MUTAGEN 219 219 D->N: NO CHANGE IN ACTIVITY.  
 SQ SEQUENCE 243 AA; 27435 MW; 65118BD9AAD609A5 CRC64;

Query Match 12.8%; Score 233.5; DB 1; Length 243;  
 Best Local Similarity 33.7%; Pred. No. 6,1e-10;  
 Matches 58; Conservative 26; Mismatches 77; Indels 11; Gaps 6;

OY 23 TNYSAKDFSGAEIYTLLEVOYGFKEARKMAAAGTSSMFLYONGSEIADGRPWVEVD 82  
 DB 81 TSPSYNKKFDCGKRSYQTYGYGLIYVNMKPAKNGIVSSFFTYTGT---DGNPWEIDI 137  
 OY 83 EVLKRNPSFOSNIITGAGAKTSEKHAVSPADQAFHYTGLEWTNRYKWTYDGOEY 142

DB 138 EFLGKDTKTVQENFYNTNGVC---NHEKIVNLGFDANASYHTTAFDQPNISIKMYVDGO-L 193  
QY 143 RTEGGQGVNLTGCTGRLFRNLWSSEA-AMVGFQFDSKLPLOFIMWVKYK 193  
DB 194 KHTATTOITQTPQPK--IMMNLNMGACVDEMIGSYN-GVTPLSRLHWRVYTK 242

RESULT 8  
YQ46\_YEAST  
ID YQ46\_YEAST STANDARD: PRT: 507 AA.  
AC P53301;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 52.8 kDa protein in BUB1-HIP1 intergenic region.  
GN YGR189C OR G7553.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C;  
RX MEDLINE=97279231; PubMed=9133739;  
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,  
RA Nombela C.;  
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm  
RL of Saccharomyces cerevisiae chromosome VII."  
RL Yeast 13:357-363(1997).  
CC -1- SIMILARITY: SOME, TO YEAST UTR2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL: 272974; CAA97215.1; -  
CC DR EMBL: X99074; CAA67525.1; -  
CC DR HSSP: P23904; IATK.  
CC SGI: S0003421; CRHL.  
CC InterPro: IPR000757; Glyco\_hydro\_16.  
CC Pfam: PF00722; Glyco\_hydro\_16; 1.  
KW Hypothetical protein.  
FT DOMAIN 63 66 POLY-SER.  
FT DOMAIN 301 310 POLY-SER.  
FT DOMAIN 345 357 POLY-SER.  
FT DOMAIN 387 391 POLY-SER.  
FT DOMAIN 467 470 POLY-SER.  
SQ SEQUENCE 507 AA; 52757 MW; 70D7B61F57AE942C CRC64;

Query Match 12.6%; Score 230; DB 1; Length 507;  
Best Local Similarity 26.9%; Pred. No. 2.7e-09;  
Matches 87; Conservative 57; Mismatches 134; Indels 46; Gaps 13;

QY 6 TAVKSAIAVAAAAALTTVNSAKDFGAEIYTLLEVOYGFKEFRMKMAASGVSSMFLY 65  
DB 70 TDLKAGETIKYSGDLSMTL-AKRYDNPGLSKSFYIMYKLEIYILKAANGTGIVSSFYIQ 128  
QY 66 ONGSEIADGRPWVEVDIEVLGNKPGSFQSNITGKAGAKTSEKHNHVPADQAEHTYG 125  
DB 129 SDDL-D-----EIDIEWVGQNTQFQSNFPSKDTTYYDREPHGVDPIDK-FHNHT 179  
QY 126 LEWTPMYVWYVDGQEVKRTGEGVSNLTGTG-----LRFNLWS---ESAA---W 171  
DB 180 LPMAMDKTMYLDGSESVR-----VLSNTSSEGYPOSPMYLMGMIVAGCDPDMAAGTIEG 233  
QY 172 V-GQPESEKLPLOFIMWVKYKYPFGQEGSDFTLDMTDNDVTFDGSRMKCGDTPFG 230  
DB 234 AGGETIYNDAPEFTMYTEKVIIVDYSTGKRYTYGDSGSG-ESTIADGSGIYRGDAQGR- 291

QY 231 NRVDLTKNIYSRGMILLALTRKG--QESFNGQVPRDEPAPQSSSAPASS--SSVPA 286  
DB 292 -----DFVLANGSGISSSTSSSTVSSASSTVSSSVSSSTVSSASSTVSSSVTS 345  
QY 287 SSSSVPASSSSAFVPPSSSSATNA 310  
DB 346 SSSSVSSSSSTS---PSSSTATSS 366

RESULT 9  
GDB\_PAEMA  
ID GDB\_PAEMA STANDARD: PRT: 237 AA.  
AC P23904;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)  
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).  
OS Paenibacillus macerans (Bacillus macerans).  
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.  
OX NCBI\_TaxID=44252;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91109712; PubMed=2274030;  
RA Borriess R., Bueltner K., Maentsaelae P.;  
RT "Structure of the beta-1,3-1,4-glucanase gene of Bacillus macerans:  
RL homologues to other beta-glucanases."  
RL Mol. Gen. Genet. 222:278-283(1990).  
RN [2]  
RP ACTIVE SITE.  
RX MEDLINE=93094208; PubMed=1360982;  
RA Hoej P.B., Condron R., Traeger J.C., McAuliffe J.C., Stone B.A.;  
RT "Identification of glutamic acid 105 at the active site of Bacillus  
RT amyloliquefaciens 1,3-1,4-beta-D-glucan 4-glucanohydrolase using  
RT epoxide-based inhibitors."  
RL J. Biol. Chem. 267:25059-25066(1992).  
RN [3]  
RP MUTAGENESIS OF GLU-128.  
RA Olsen O.;  
RL Thesis (1990), University of Aarhus, Denmark.  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=93281743; PubMed=8099449;  
RA Keitel T., Simon O., Borriess R., Heinemann U.;  
RT "Molecular and active-site structure of a Bacillus 1,3-1,4-beta-  
RL glucanase."  
RL Proc. Natl. Acad. Sci. U.S.A. 90:5287-5291(1993).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
RX MEDLINE=96028129; PubMed=7588726;  
RA Hahn M., Keitel T., Heinemann U.;  
RT "Crystal and molecular structure at 0.16-nm resolution of the hybrid  
RT Bacillus endo-1,3-1,4-beta-D-glucan 4-glucanohydrolase HAl6-M.";  
RL Eur. J. Biochem. 232:849-858(1995).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages  
CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE  
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY  
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL: X55959; CAA39426.1; -  
CC PIR: S11927; S11927.  
CC PDB: 2AVH; 31-MAR-95.  
CC PDB: 1BYH; 31-OCT-93.  
CC PDB: 1CPM; 22-JUN-94.

Rak B.;  
RT "Licit, a Bacillus subtilis transcriptional antiterminator protein of  
RT the BglG family";  
RT J. Bacteriol. 178:1971-1979(1996).  
RN [6]  
RP SEQUENCE OF 29-63.  
RA Yuki T., Tezuka H., Yabuuchi S.;  
RT "Purification and some properties of two enzymes from a beta-glucanase  
RT producing strain, Bacillus subtilis HL-25";  
RT Agric. Biol. Chem. 53:2341-2346(1989).  
RL  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages  
CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE  
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.  
CC -1- MISCELLANEOUS: THE SEQUENCE OF STRAIN 168 IS SHOWN.  
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X00754; CAA25328.1; -  
CC EMBL; D00518; BAA00405.1; -  
CC EMBL; D83026; BAA11697.1; -  
CC EMBL; Z46862; CAA86922.1; -  
CC EMBL; Z28340; CAA82195.1; -  
CC EMBL; Z29124; CAA15943.1; -  
CC PIR; A22914; LMB5.  
CC PIR; J00110; J00110.  
CC HSSP; P27051; 1GBG.  
DR Subtilisin; BG10476; bg1s.  
DR InterPro; IPR000757; Glyco\_hydro\_16.  
DR Pfam; PF00722; Glyco\_hydro\_16; 1.  
DR PRINTS; PR00737; GLHYDRLASE16.  
DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
DR KMW Hydrolyase: Glycosidase; Signal: Complete proteome.  
FT SIGNAL 1 28  
FT CHAIN 1 29  
FT MOD\_RES 29 29 BETA-GLUCANASE.  
FT ACT\_SITE 133 133 PYRROLIDONE CARBOXYLIC ACID.  
FT ACT\_SITE 137 137 NUCLEOPHILE (BY SIMILARITY).  
FT DISULFID 60 89 PROTON DONOR (BY SIMILARITY).  
FT VARIANT 24 24 A -> S (IN STRAIN HL-25).  
FT VARIANT 83 83 A -> S (IN STRAIN HL-25).  
FT VARIANT 204 204 P -> L (IN STRAIN C120).  
FT VARIANT 204 204 P -> L (IN STRAIN C120).  
SQ SEQUENCE 242 AA; 27268 MW; 4595BDEA70F22B29 CRC64;  
  
Query Match 13.6%; Score 248; DB 1; Length 242;  
Best Local Similarity 36.0%; Pred. No. 5.6e-11;  
Matches 63; Conservative 25; Mismatches 75; Indels 12; Gaps 7;  
  
OY 20 ALTVNSAKDFGAEIYTLDEYQKFEARMKMAAGSTVSMPLYONGSEIADGRPWVEV 79  
DB 78 ALTSFAVYK - FDCGNSRSQVGTGYGLYEVKRMKPAKNTGIVSSFFTYTGFT ---DGPWDE 133  
OY 80 VDIEVLKNGPSFGSNIITGKAGAKTSEKHVAVSPADQAFHTYGLEMTPNVYRWTV 139  
DB 134 IDIEFLGDTKTVQNFNYTNGAG ---NHEKIVDLGFDANAVHTAFAFDMPQNSIKWYVDG 190  
OY 140 QEVRTREGGVSNLTGQGLRNLMSSESA -AWVGQFDESKLPLOFINWVYVK 193  
DB 191 Q-LKHTATNOLPTTPCK -IMMNLNMGTVGVDMLGSLYNGVN -PLVAHYDWMRYRK 241  
RESULT 4  
GUB\_BACAM STANDARD; PRT; 239 AA.  
AC P07980;

DT 01-AUG-1988 (Rel. 03, Created)  
DT 01-AUG-1988 (Rel. 03, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)  
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).  
GN BGLA.  
OS Bacillus amyloliquefaciens.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1390;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BE 20/78;  
RX MEDLINE-87192007; PubMed-3106158;  
RA Hofemeister J., Kuriz A., Boriss R., Knowles J.;  
RT "The beta-glucanase gene from Bacillus amyloliquefaciens shows  
RT extensive homology with that of Bacillus subtilis";  
RL Gene 49:177-187(1985).  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glucosidic linkages  
CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE  
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.  
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M15674; AAA87323.1; -  
CC PIR; A29091; A29091.  
CC HSSP; P27051; 1GBG.  
DR InterPro; IPR000757; Glyco\_hydro\_16.  
DR Pfam; PF00722; Glyco\_hydro\_16; 1.  
DR PRINTS; PR00737; GLHYDRLASE16.  
DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
DR KMW Hydrolyase: Glycosidase; Signal.  
FT SIGNAL 1 25  
FT CHAIN 1 26  
FT ACT\_SITE 134 134 BETA-GLUCANASE.  
FT DISULFID 57 86 NUCLEOPHILE (BY SIMILARITY).  
SQ SEQUENCE 239 AA; 26928 MW; A76A64268A7AA0B CRC64;  
  
Query Match 13.3%; Score 242.5; DB 1; Length 239;  
Best Local Similarity 34.3%; Pred. No. 1.4e-10;  
Matches 59; Conservative 27; Mismatches 75; Indels 11; Gaps 6;  
  
OY 23 TVNSAKDFGAEIYTLDEYQKFEARMKMAAGSTVSMPLYONGSEIADGRPWVEV 82  
DB 77 TSPSYNKFDCGNSRSQVGTGYGLYEVKRMKPAKNTGIVSSFFTYTGFT ---GTPWDEIDI 133  
OY 83 EVLKNGPSFGSNIITGKAGAKTSEKHVAVSPADQAFHTYGLEMTPNVYRWTV 142  
DB 134 EFLGDTKTVQNFNYTNGAG ---NHEKIVDLGFDANAVHTAFAFDMPQNSIKWYVDG 189  
OY 143 KTEGGOVSNLTGQGLRNLMSSESA -AWVGQFDESKLPLOFINWVYVK 193  
DB 190 KHTATTOIPAPDGK -IMMNLNMGTVGVDMLGSLYNGVN -PIVAHYDWMRYRK 238  
RESULT 5  
XYND\_RUMFL STANDARD; PRT; 802 AA.  
AC Q53317;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Xylanase/beta-glucanase precursor (includes: Endo-1,4-beta-xylanase  
DE (EC 3.2.1.8) (Xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)  
DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase)).  
GN XYND.

OS Ruminococcus flavefaciens.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;  
 OC Ruminococcus.  
 OX NCBI\_TaxID=1265;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-17;  
 RX MEDLINE=93259938; PubMed=8491715;  
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;  
 RT "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-  
 RT glucanase domains, encoded by the xynD gene of Ruminococcus  
 RT flavefaciens";  
 RL J. Bacteriol. 175:2943-2951(1993).  
 CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-  
 CC BETA-1,3-1,4 GLUCANASE ACTIVITIES.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic  
 CC linkages in xylans.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages  
 CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
 CC -1- PATHWAY: xylan degradation.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY  
 CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF  
 CC GLYCOSYL HYDROLASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: S61204; AAB2620.1; -  
 CC DR HSSP: P23904; IAAK.  
 DR InterPro: IPR003305; CBM\_CenC.  
 DR InterPro: IPR001137; GH\_11.  
 DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00457; Glyco\_hydro\_11; 1.  
 DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
 DR Pfam: PF02018; CBM\_4\_9; 1.  
 DR PRINTS: PR00911; GLHYDRLASE1.  
 DR PRINTS: PR00737; GLHYDRLASE16.  
 DR PROSITE: PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE: PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR Xylan degradation: Hydrolase; Glycosidase; Signal;  
 DR KM Multifunctional enzyme.  
 FT SIGNAL 1 31  
 FT CHAIN 32 802  
 FT DOMAIN 32 244  
 FT DOMAIN 245 523  
 FT DOMAIN 524 555  
 FT DOMAIN 556 802  
 FT ACT\_SITE 124 124  
 FT ACT\_SITE 226 226  
 FT ACT\_SITE 684 684  
 FT DOMAIN 524 529  
 FT DOMAIN 532 543  
 FT DOMAIN 546 553  
 FT SEQUENCE 802 AA: 89091 MW: 286046896472844F CRC64;  
 Query Match 13.3%; Score 242.5; DB 1; Length 802;  
 Best Local Similarity 34.9%; Pred. No. 6e-10;  
 Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

OY 30 FSGAEIYIEEVOYKFEKFRMKMAASGVSSFLYQNSGLADGRPWEVDIEVLGNP 89  
 Db 634 YSGGEFRTNNFYHYGYECESQAMKNDGVSSFFETGTPS---DDNPMEIDIEILGKNT 690  
 OY 90 GSFOSNIITGKGAQKADQTSKHNHVAAPDAQFHTGLTEPRNVYRTVAGQVYKRTESQ 149  
 Db 691 TGVQFRVYITNGGCKH---EKLTDGFDSSSEAVHTYGFDPNTIAYVDGREVYRA--- 743

RESULT 6  
 GUB\_CLOTM  
 ID GUB\_CLOTM STANDARD; PRT; 334 AA.  
 AC P29716; P37074;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)  
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase) (Laminarinase).  
 GN LICB OR LAM1.  
 OS Clostridium thermocellum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 27405 / DSM 1237;  
 RX MEDLINE=92155194; PubMed=1740123;  
 RA Schramm S., Schwarz W.H., Staudenbauer W.L.;  
 RT "Structure of the Clostridium thermocellum gene licB and the encoded  
 RT beta-1,3-1,4-glucanase. A catalytic region homologous to Bacillus  
 RT cellulases joined to the reiterated domain of clostridial  
 RT cellulases";  
 RL Eur. J. Biochem. 204:13-19(1992).  
 RL [2]  
 RP PRELIMINARY SEQUENCE FROM N.A.  
 RC STRAIN-E7;  
 RX MEDLINE=92095946; PubMed=1755832;  
 RA Zverlov V.V., Laptev D.A., Tishkov V.I., Velikodvorska G.A.;  
 RT "Nucleotide sequence of the Clostridium thermocellum laminarinase  
 RT gene";  
 RL Biochem. Biophys. Res. Commun. 181:507-512(1991).  
 RL [3]  
 RP REVISIONS.  
 RA Zverlov V.V.;  
 RL Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP CHARACTERIZATION.  
 RC STRAIN-E7;  
 RA Zverlov V.V., Velikodvorska G.A.;  
 RT "Cloning the Clostridium thermocellum thermostable laminarinase gene  
 RT in Escherichia coli; the properties of the enzyme thus produced.";  
 RL Biotechnol. Lett. 12:811-816(1990).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages  
 CC in beta-D-glucans containing 1,3- and 1,4-bonds.  
 CC -1- SUBUNIT: MAY FORM PART OF A MULTENZYME COMPLEX (CELLULOSE).  
 CC -1- DOMAIN: A 24 RESIDUE DOMAIN IS REPEATED TWICE IN THIS ENZYME AS  
 CC WELL AS IN OTHER C.THERMOCELLUM CELLULOSE ENZYMES. THIS DOMAIN  
 CC MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X63355; CAA44959.1; -  
 DR EMBL: X58392; CAA41281.1; -  
 DR PIR: S23498; S23498.  
 DR PIR: JS0611; JS0611.  
 DR PIR: S18726; S18726.  
 DR HSSP: P23904; IAAK.  
 DR InterPro: IPR002105; Dockerin\_1.  
 DR InterPro: IPR002048; EF-hand.

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:50:42 : Search time 125.774 Seconds  
(without alignments)  
1368.672 Million cell updates/sec

Title: US-09-654-652A-2

Perfect score: 1439  
Sequence: 1 MVSADPFGAELTYEEVQY.....PNSSVYDLAALHHHHHH 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/1/paa/PCRTUS\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep: \*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep: \*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep: \*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep: \*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep: \*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep: \*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep: \*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep: \*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep: \*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep: \*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep: \*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep: \*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep: \*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep: \*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep: \*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep: \*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep: \*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep: \*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep: \*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep: \*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep: \*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1439	100.0	267	20	US-09-654-652A-2
2	1340	93.1	349	20	US-09-654-652A-3
3	1340	93.1	349	21	US-09-791-537-30441
4	1333	92.6	248	20	US-09-654-652A-1
5	272.5	18.9	214	21	US-09-791-537-145019
6	266.5	18.5	259	21	US-09-791-537-106873

7	265.5	18.5	252	21	US-09-791-537-32335
8	257.5	17.9	308	18	US-09-463-862A-1
9	256.5	17.8	208	21	US-09-791-537-6910
10	250.5	17.4	394	21	US-09-791-537-132244
11	247.5	17.2	214	21	US-09-791-537-117361
12	247.5	17.2	276	21	US-09-791-537-31771
13	246.5	17.1	242	21	US-09-791-537-27063
14	245.5	17.1	242	21	US-09-791-537-73450
15	245.5	17.1	242	21	US-09-791-537-79889
16	245	17.0	239	21	US-09-791-537-103347
17	244	17.0	239	21	US-09-791-537-103344
18	242.5	16.9	802	21	US-09-791-537-103320
19	242.5	16.9	802	21	US-09-791-537-116925
20	238.5	16.6	239	21	US-09-791-537-30436
21	237	16.5	334	21	US-09-791-537-48909
22	235.5	16.4	214	21	US-09-791-537-88552
23	234.5	16.3	239	21	US-09-791-537-18679
24	233	16.2	239	21	US-09-791-537-103817
25	231.5	16.1	320	21	US-09-791-537-67048
26	230	16.0	237	21	US-09-791-537-72848
27	229.5	15.9	243	21	US-09-791-537-30440
28	226	15.7	212	21	US-09-791-537-88873
29	226	15.7	214	21	US-09-791-537-88270
30	226	15.7	237	21	US-09-791-537-103320
31	226	15.7	237	21	US-09-791-537-106876
32	226	15.7	239	21	US-09-791-537-103345
33	226	15.7	239	21	US-09-791-537-103350
34	226	15.7	239	21	US-09-791-537-103355
35	226	15.7	239	21	US-09-791-537-103357
36	226	15.7	239	21	US-09-791-537-103819
37	226	15.7	240	21	US-09-791-537-64827
38	226	15.7	240	21	US-09-791-537-79887
39	225	15.6	238	21	US-09-791-537-37095
40	222.5	15.4	242	21	US-09-791-537-37095
41	222	15.4	237	21	US-09-791-537-3485
42	209.5	14.9	160	21	US-09-706-729-14
43	209.5	14.6	504	27	US-60-385-568-329
44	209.5	14.6	526	25	US-10-179-131-6380
45	208.5	14.5	526	16	US-09-248-796-14807

ALIGNMENTS

RESULT 1  
US-09-654-652A-2  
Sequence 2, Application US/09654652A  
GENERAL INFORMATION:  
APPLICANT: SHYUR, LIE-PEN  
APPLICANT: CHEN, JUI-LIN  
APPLICANT: YANG, NING-SUN  
TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLAST SUCCINOGENES 1, 3-1,  
TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMAIC ACTIVITY AND  
FILE REFERENCE: 4910-8  
CURRENT APPLICATION NUMBER: US/09/654,652A  
CURRENT FILING DATE: 2000-09-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme  
OTHER INFORMATION: with enhanced activity and thermal stability  
US-09-654-652A-2

Query Match 100.0%; Score 1439; DB 20; Length 267;  
Best Local Similarity 100.0%; Pred. No. 5.1e-138;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MVSADPFGAELTYEEVQYGRFEARKKMAASGTSSMFLYNGSEINDGRWVEVDIE 60

```
|||||
Db 1 MVSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 60
QY 61 VLGNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 120
   |||||||
Db 61 VLGNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 120
QY 121 KTEGGQVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGGS 180
   |||||||
Db 121 KTEGGQVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGGS 180
QY 181 DFTLDWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 240
   |||||||
Db 181 DFTLDWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 240
QY 241 PRDEPAPNSSVDKLAALAEHHNNH 267
   |||||||
Db 241 PRDEPAPNSSVDKLAALAEHHNNH 267
```

## RESULT 2

```
US-09-654-652a-3
; Sequence 3. Application US/09654652a
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLASTER SUCCINOGENES 1, 3-1,
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMATIC ACTIVITY AND
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654, 652a
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-654-652a-3
```

```
Query Match          93.1% Score 1340; DB 20; Length 349;
Best Local Similarity 99.6% Pred. No. 1e-127;
Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 VSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 61
   |||||||
Db 25 VSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 61
QY 62 LGRNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 121
   |||||||
Db 85 LGRNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 144
QY 122 TEGGOVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGSD 181
   |||||||
Db 145 TEGGOVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGSD 204
QY 182 FTLDMWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 241
   |||||||
Db 205 FTLDMWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 264
QY 242 RDDEPAPNSS 252
   |||||||
Db 265 RDDEPAPNSS 275
```

## RESULT 3

```
US-09-791-537-30441
; Sequence 30441. Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
```

```
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791, 537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30441
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Fibrobacter succinogenes
US-09-791-537-30441
```

```
Query Match          93.1% Score 1340; DB 21; Length 349;
Best Local Similarity 100.0% Pred. No. 3.2e-127;
Matches 248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 VSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 61
   |||||||
Db 25 VSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 61
QY 62 LGRNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 121
   |||||||
Db 85 LGRNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 144
QY 122 TEGGOVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGSD 181
   |||||||
Db 145 TEGGOVSNLTGTGGLRFLNLMSSSAAMVGOFPDESKLPLEOFIMVVKYKTPGQEGGSD 204
QY 182 FTLDMWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 241
   |||||||
Db 205 FTLDMWTDNFTFDGSRMGKGDWTFDGNRYDLTDKNIYSRDGMILALTRKGQSFNGQV 264
QY 242 RDDEPAPNSS 252
   |||||||
Db 265 RDDEPAPNSS 275
```

## RESULT 4

```
US-09-654-652a-1
; Sequence 1. Application US/09654652a
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLASTER SUCCINOGENES 1, 3-1,
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654, 652a
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATUERS:
; OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
; OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652a-1
```

```
Query Match          92.6% Score 1333; DB 20; Length 248;
Best Local Similarity 100.0% Pred. No. 3.2e-127;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MVSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 60
   |||||||
Db 1 MVSAKDFSGAELTYLLEVOYGFKFEARMKMAASGTVSSMFLYONGSEIADGRPWVEVDIE 60
QY 61 VLGNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 120
   |||||||
Db 61 VLGNPGSFOSNIITTKAGAKOCTSEKHNHVAVSPADDAFHHTYGLEMTPNYVRYMTVDGQEVK 120
```





```

; LENGTH: 308
; TYPE: PRT
; ORGANISM: Bacillus alkalophilus DSM 9956
US-09-463-862A-1

```

Query Match	17.9%	Score 257.5	DB 18	Length 308
Best Local Similarity	36.5%	Pred. No. 4.4e-17		
Matches 62; Conservative		22; Mismatches 73;	Indels 13;	Gaps 6

[illegible]

RESULT 9  
US-09-791-537-6910

Query Match	17.8%	Score	256.5	DB	21	Length	208
Best Local Similarity	33.8%	Pred. NO.	3.2e-17				
Matches	77	Conservative	29	Mismatches	95	Indels	27
						Gaps	13

QY 7 FSGSLVLYLEEVQXGKFFAPRRKKMAAASQTSVMPLYONGSEIADGRPMVEVDIEVLKNP 66  
 Db 1 FDCAEYRSTNITNYIGLYEVSMKKPAKNTGYSSFTTYGP---AHGTOMDEIDIEFLKDT 57  
 QY 67 GSFOSNITTKAGAGAKTSEKIHNAVSPADQAFHYLYGLEMTPNVRYMTVDGQEVKRTGGQ 126  
 Db 58 TKVQENYNTNGVGH---EKVYISLGFPASKKEFHNYARNDMQPGYIKMYVLD---A 110  
 QY 127 VSNLTGTGQ-LRENLWSEESA-AVWQGFDESKLPFLQFINWKKYKYKTPQGECEGSDFTL 184  
 Db 111 TANIPSTPGKIMMLMNGTGVDDMLGYSNGCAN-PLVLEYDVA---KYT-----SNGSYF- 160  
 QY 185 DWTDJDFDTFDDGSRNGKGDWMTFDGNRYDLT--DKNI-VSRQMLTLALT 229  
 Db 161 -WEPK-STYFNPSJTEWKADGYSNGGVFNCMTARANNVNTNIGKLIKGLIT 206

```

RESULT 10
US-09-791-537-132244
: Sequence 132244, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210

```

```

; CURRENT APPLICATION NUMBER: US/09/791,537
;
; CURRENT FILING DATE: 2001-02-22
;
; NUMBER OF SEQ ID NOS: 153055
;
; SOFTWARE: PatentIn version 3.0

```

Query Match:	17.4%	Score 250.5:	DB 21:	Length 394:
Best Local Similarity:	29.68%	Pred. No. 3.3e-16:		
Matches 68;	Conservative 33;	Mismatches 90;	Indels 39;	Gaps 10

```

0Y      7 FSGALYLYLEEYQKGFAPARKMAAASCTYSMPLYONSGEADGRPMVDEIVLCKNP 66
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      1 FDCATYRSTINITYGIVLYSVKPKAKNTGIVSFFTYGP---AHGTOMDEIDIEFLCKDT 57

0Y      67 GSFOSNIITTKAGAQOKTSEKHHAVSPADAQFHTYTGLEMTPNVRYMTVDGQEVARKTEGCO 126
      58 TKVQENYNTNGVGH---EKVYISLGFQDSKGFHYARDMQPGYIKWYVDG-VLTKT---A 110
Db

0Y      127 VSNLTGTGG-LRFNIMSSESA-AWVGQFDESKLPLFQFINWKVY-----YKTPGQ 175
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      111 TANIPSTPGKIMMNLMTNGTGVDDMLGYSYNGAN-PLVLEYDMVKYTSNASTDYQMOMTWDG 169

0Y      176 G-----EGGSDFTLDMWTDNFDTEGDSRWGCGD-----WTEDGN 208
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      170 GIVNAVNSGGNYSYVMSNTGNEFVVGKMTTGSPEFTIINAGYMAADGN 219

```

```

RESULT 11
US-09-791-537-117361
: Sequence 117361 Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 117361
: LENGTH: 214
: TYPE: PRT
: ORGANISM: pdb 1AUKA
US-09-791-537-117361

```

Query Match	17.28;	Score 247.5;	DB 21;	Length 214;
Best Local Similarity	32.98;	Pred. No. 2.8e-16;		
Matches 69; Conservative	32;	Mismatches 84;	Indels 25;	Gaps 12

```

OY      33  SCGYSSMFLYONGSEIADGRPMVEDEIVLEKKNPSSFOSNITTCKAQAQTSKHNHVP  92
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2  TGYSSFFTYTGP---AHGTOMDEIDIEFLCKDTTKQFNYITNGVGH---EKVYSLGF  55

OY      93  AADQAFHTYGLEMTPNVRYRMTVDGQEVKRTGCGGVSNLTGTG--LRFNLWSSESA--AWG  150
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      56  DASKRFHTYARDDMQPGYIKMWVDG--VLKHT---ATANIPTPCKIMMNLNNGTVDDMLG  111

OY      151  QFDESKLPLEFQIMWVKYKYTTPGQGEBCGSDFTLDMTDNFPFEDGSMWKGDMTFDGNRY  210
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      112  SYNGAN--PLYAEYDMV---KYTSMQ--TGGSPE-----EPFNSYSGTWEKADGVSNGVF  161

OY      211  DLT--DKNI--YSRDOMLILATIRRGQSESN  237
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      162  NCTWRANVNYFTNDGCKLTKGLTSSANYMKFD  191

```

## RESULT 12



Thu Jan 9 12:09:15 2003

us-09-654-652a-2.rapm

Page 6

Search completed: January 9, 2003, 12:13:28  
Job time : 127.774 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 31.103 Seconds  
(without alignments)  
2312.009 Million cell updates/sec

Title: US-09-654-652a-3  
Perfect score: 1824  
Sequence: 1 NNIKKTAVKSALVAIAAAAAA.....AKGAKVNPNGKRYRVNFEH 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-virus:\*  
13: SP-vertebrate:\*  
14: SP-unclassified:\*  
15: SP-virus:\*  
16: SP-bacteriap:\*  
17: SP-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	268.5	14.7	851	16	Q9K7X5
2	254.5	14.0	256	2	Q9APD8
3	247.5	13.6	276	2	Q45648
4	246.5	13.5	214	2	Q93GE8
5	245.5	13.5	214	2	Q93GE7
6	242.5	13.3	802	2	Q9S310
7	239.5	13.1	214	2	Q93GE6
8	238.5	13.1	239	2	Q45691
9	236.5	13.0	242	2	Q8RMP0
10	232.5	12.7	212	2	Q9FDC9
11	230	12.6	237	2	Q07856
12	223	12.2	205	2	Q93U11
13	221	12.1	246	16	Q97FD3
14	206	11.3	289	16	Q8U8N5
15	200	11.0	293	16	Q98C78
16	190.5	10.4	245	3	O14412

17	190.5	10.4	302	16	Q9K7X6	Q9K7X6 bacillus ha
18	185	10.1	282	10	Q9F131	Q9F131 arabidopsis
19	178	9.8	302	3	O42800	O42800 aspergillus
20	173	9.5	282	10	O80803	O80803 arabidopsis
21	173	9.5	642	16	Q9WXX1	Q9WXX1 thermotoga
22	172.5	9.5	277	10	Q9W0D1	Q9W0D1 arabidopsis
23	171	9.4	646	2	O60039	O60039 thermotoga
24	170.5	9.3	282	10	Q9W0D2	Q9W0D2 arabidopsis
25	169.5	9.3	277	10	Q9A4A9	Q9A4A9 arabidopsis
26	164	9.0	286	10	Q9B910	Q9B910 arabidopsis
27	161	8.8	284	10	Q9FKL8	Q9FKL8 arabidopsis
28	160	8.8	286	10	O8W4M6	O8W4M6 arabidopsis
29	159.5	8.7	292	10	Q9ZS04	Q9ZS04 arabidopsis
30	159.5	8.7	306	2	O51333	O51333 oerskovia x
31	157.5	8.6	92	3	Q9HCU1	Q9HCU1 aspergillus
32	157.5	8.6	94	3	Q9PA20	Q9PA20 erlichoderma
33	157.5	8.6	269	10	Q9J148	Q9J148 arabidopsis
34	156	8.6	292	10	O9ZRV1	O9ZRV1 faqus sylvia
35	155	8.5	845	2	Q9KWF3	Q9KWF3 clostridium
36	155	8.5	845	16	O8XNF8	O8XNF8 clostridium
37	154.5	8.5	277	10	O38907	O38907 arabidopsis
38	154.5	8.5	284	10	Q9SEB0	Q9SEB0 arabidopsis
39	153	8.4	163	2	Q930I2	Q930I2 uncultured
40	149	8.2	289	10	P93669	P93669 hordeum vul
41	146.5	8.0	284	10	O38857	O38857 arabidopsis
42	145.5	8.0	285	10	O9FKL9	O9FKL9 arabidopsis
43	145.5	8.0	310	10	O9X1J7	O9X1J7 arabidopsis
44	145	7.9	280	10	O94910	O94910 festuca pra
45	145	7.9	305	10	O9ZV40	O9ZV40 arabidopsis

#### ALIGNMENTS

RESULT 1  
ID Q9K7X5 PRELIMINARY: PRT: 851 AA.  
AC Q9K7X5:  
DT 01-OCT-2000 (TREMBL:rel. 15, Created)  
DT 01-OCT-2000 (TREMBL:rel. 15, Last sequence update)  
DE 01-DEC-2001 (TREMBL:rel. 19, Last annotation update)  
DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).  
GN BGLS OR BH3232.  
OS Bacillus halodurans.  
OC Bacteria: Firmicutes: Bacillus/Clostridium group: Bacillales:  
OC Bacillaceae: Bacillus.  
OC NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512562; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR HMBL: AP001518; BAB06951.1; -.  
DR HSSP: P23904; IAJK.  
DR InterPro: IPR00757; Glyco\_hydro\_16.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Pfam: PF00722; Glyco\_hydro\_16; 3.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR PRINTS: PR00737; GLHYDRLASE16.  
DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 3.  
KW Hydrolyase; Glycosidase; Complete proteome.  
SQ SEQUENCE 851 AA: 98142 MW; C1CA4F0F158400285 CRC64;

Query Match 14.7%; Score 268.5; DB 16; Length 851;  
Best Local Similarity 29.4%; Pred. No. 7.2e-10;  
Matches 83; Conservative 32; Mismatches 112; Indels 55; Gaps 12;

QY 21 LTTNVAKDSGAEIVLTLEVGKGFARMKMAASGVSMFLYQNGSFRIADGRWVEV 80

Db 93 LENDQAPPKAGELRTNDYQYGLFEVSMKPAKVECTVSSFFTY-TGEWDWGDDEI 151  
QY 81 DIEVLKNGPSFOSNIITGKAGAKTSEKHAAVSPADQAFHTYGLFWTNNYRWTVDQ 140  
Db 152 DIEFLGKOTTRIOENFTYNGVCG---NEHYTDGFEDSSEFNNTYAFEMREDSTWYNGE 208  
QY 141 EVKRTGEGOVSNLTGT-QGLRFNLMSSEA-AWVGQFDESKLPLFOFINWKKYKYP-- 196  
Db 209 AVHTA---TENIPQFQKLMNMLMPGVGVDEMTGVDFGNFTPLHADYENV---RYTPLE 261  
QY 197 --GQEGEGSFTLDWT-----DNFTFDGSRWG-KGDWTFDG 230  
Db 262 ALDEESGDNEEPEVEEVEEPEADEVSVRIGSAIYETFTYFNEDIMSIAGWT-NG 320  
QY 231 NRYDLT--DKNIYSRDCMLLALTRKQOESFNQVPRADPEA 270  
Db 321 QMENATWYDSQITFSNCFMFAIDKE-----DDEEA 351

## RESULT 2

Q9APD8 PRELIMINARY: PRT: 256 AA.  
AC Q9APD8; 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Endo-beta-1,3,1,4-glucanase.  
GN BGLBCL  
OS Bacillus circulans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC21367;  
RA Lee D.-S., Kim J.-Y., Kim H.-B.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306531; AAG53947.1; -  
DR HSSP: P23904; IAJK.  
DR InterPro: IPR000757; Glyco\_hydro\_16.  
DR Pfam: PR00722; Glyco\_hydro\_16.  
DR PRINTS: PR00737; GLYCDRLASE16.  
DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
SQ SEQUENCE 256 AA: 27604 MW: 989P50C09F6E4531 CRC64;

Query Match 14.0%; Score 254.5; DB 2; Length 256;

Best Local Similarity 34.6%; Pred. No. 1.3e-09;

Matches 62; Conservative 29; Mismatches 71; Indels 17; Gaps 7;

QY 23 TTVSANDFSGAELTYLLEVOYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82  
Db 88 SNGSGRPYASAEVATQKGYGRVEARIKAKGTGLVTSIFTY---SGAARPGISNDEIDI 144  
QY 83 EVLGNPGSFOSNIITGKAGAKTSEKHAAVSPADQAFHTYGLFWTNNYRWTVDQ 140  
Db 145 EILGKPTTMMETNYFTNGVGHST-----VIDGFDAISLDHFYAFEMSPTSIKWYVDGR 199  
QY 141 EVKRTGEGOVSNLTGT-QGLRFNLMSSEA-AWVGQFDESKLPLFOFINWKKYKYP 196  
Db 200 LVH-TETGSRGPLPTSPGYIMVLMWSGAPAEIWTGKFTYPCGPIRAYDWM---KETP 254

## RESULT 3

Q45648 PRELIMINARY: PRT: 276 AA.  
AC Q45648; 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Lichenase precursor.  
GN BGLI  
OS Bacillus sp.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1409;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=137;  
RA MEDLINE=94288605; PubMed+7517127;  
RA Tabernero C., Coll P.M., Fernandez-Abalos J.M., Perez P.,  
RA Santamaria R.I.;  
RT "Cloning and DNA sequencing of bgaA, a gene encoding an endo-beta-1,3-  
1,4-glucanase, from an alkalophilic Bacillus strain (N137).";  
RL Appl. Environ. Microbiol. 60:1213-1220(1994).  
DR EMBL: Z12151; CAA78135.1; -  
DR HSSP: P23904; IAJK.  
DR InterPro: IPR000757; Glyco\_hydro\_16.  
DR Pfam: PR00722; Glyco\_hydro\_16; 1.  
DR PRINTS: PR00737; GLYCDRLASE16.  
DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
FT SIGNAL.  
KW Signal.  
FT CHAIN 1 31 POTENTIAL.  
FT CHAIN 32 276 LICHENASE.  
SQ SEQUENCE 276 AA: 31770 MW: D047F8A34CA9EBE2 CRC64;

Query Match 13.6%; Score 247.5; DB 2; Length 276;

Best Local Similarity 36.1%; Pred. No. 4.3e-09;

Matches 61; Conservative 20; Mismatches 75; Indels 13; Gaps 6;

QY 30 FSGAELTYLLEVOYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVLKRP 89  
Db 96 YKAGELRTNDYHYGLFEVSMKPAKSTGTSSFFTY-TGPMWENDMPDEIDIEFLGKDT 154  
QY 90 GSFOSNIITGKAGAKTSEKHAAVSPADQAFHTYGLFWTNNYRWTVGOEVRKTEGQ 149  
Db 155 TKIQFNFTYNGVCG---NEHYHGLGFDADDFNTYAFEMRPSIRKFPVAGELVHTA----- 207  
QY 150 VSNLTGT-QGLRFNLMSSEA-AWVGQFDESKLPLFOFINWKKYKYP 196  
Db 208 TENIPQFQKLMNMLMPGVGVDEMTGVDFGNFTPLHADYENV---KYTP 253

## RESULT 4

Q93GE8 PRELIMINARY: PRT: 214 AA.  
AC Q93GE8; 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Endo 1-3,1,4-beta-glucanase (Fragment).  
OS uncultured bacterium.  
OC Bacteria; environmental samples.  
OX NCBI\_TaxID=77133;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Rodriguez V., Mellado R.P.;  
RA "Novel Lichenases from soil."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF254957; AAK50610.2; -  
DR InterPro: IPR000757; Glyco\_hydro\_16.  
DR Pfam: PR00722; Glyco\_hydro\_16; 1.  
DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; UNKNOWN\_1.  
FT NON\_TER 1 1  
SQ SEQUENCE 214 AA: 24159 MW: 3DB186D3B9D91CF CRC64;

Query Match 13.5%; Score 246.5; DB 2; Length 214;

Best Local Similarity 35.5%; Pred. No. 3.6e-09;

Matches 61; Conservative 25; Mismatches 75; Indels 11; Gaps 6;

QY 23 TTVSANDFSGAELTYLLEVOYGFKEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82  
Db 52 TSPSYNKPFCGGRNSQVQTYGYGRVEARIKAKGTGLVTSIFTY---GTPWDEIDI 108  
QY 83 EVLGNPGSFOSNIITGKAGAKTSEKHAAVSPADQAFHTYGLFWTNNYRWTVGOGEV 142

```
Db 109 EFLGKDTTKVQENYYTNGAG---NHEKLDLGFDAANAHYTFADWQPSIKWYVDGQ-L 164
OY 143 RTEGGQVSNLTGQGLRFNLMSSESA-AWGOFPDESKLPLOFINWVYVK 193
Db 165 KHTATTQITPAAGK--IMNMNMGTCVDDMLGSYGNV-PLTAHIDWVAYTK 213

RESULT 5
O93GE7 PRELIMINARY: PRT: 214 AA.
AC 093GE7:
Db 01-DEC-2001 (TREMBLrel. 19, Created)
Dr 01-DEC-2001 (TREMBLrel. 19, last sequence update)
Dr 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Endo 1-3,1-4-beta-glucanase (Fragment).
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxId=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Mellado R.P.;
RT "Novel lichenases from soil."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF254959; AAK50612.2;
DR InterPro: IPR000757; Glyco_hydro.16.
DR Pfam: PF00722; Glyco_hydro.16; 1.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS01034; GLYCOSTYL_HYDROL_F16; UNKNOWN_1.
DR NON_TER 1
FT SEQUENCE 214 AA: 24175 MW; 3DB192D2F6B4CAA3 CRC64;

Query Match 13.5%; Score 245.5; DB 2; Length 214;
Best Local Similarity 35.5%; Pred. No. 4.2e-09;
Matches 61; Conservative 25; Mismatches 75; Indels 11; Gaps 6;

OY 23 TNVSAKDFSGAELTYLEVOYGFKFEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82
Db 52 TSPSYNKFPCGERSVQVTGYGLYEVRMKPAKNTGIVSSFFTYTGPT--GTPMDEIDI 108
OY 83 EVLGNKPGSFQSNITTKGKGAOKTSEKHNAVSPADQAEHTYGLETPVYVNTVDGQEV 142
Db 109 EFLGKDTTKVQENYYTNGAG---NHEKLDLGFDAANAHYTFADWQPSIKWYVDGQ-L 164
OY 143 RTEGGQVSNLTGQGLRFNLMSSESA-AWGOFPDESKLPLOFINWVYVK 193
Db 165 KHTATTQITPAAGK--IMNMNMGTCVDDMLGSYGNV-PLTAHIDWVAYTK 213

RESULT 6
O9S310 PRELIMINARY: PRT: 802 AA.
AC 09S310:
Db 01-MAY-2000 (TREMBLrel. 13, Created)
Dr 01-DEC-2001 (TREMBLrel. 19, last sequence update)
Dr 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.
GN XYND.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Lachnospiraceae; Ruminococcus.
OX NCBI_TaxId=1265;
RN [1]
RP SEQUENCE FROM N.A.
RA Autila V., Martin J.C., Munro C.A., Mercer D.K., Flint H.J.;
RT "Organisation and strain distribution of genes responsible for the
RT utilization of xylans by the rumen cellulolytic bacterium Ruminococcus
RT flavefaciens 17."
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ132472; CAB51934.1;
DR HSPF: P23904; IATK.
DR InterPro: IPR003305; CBM_CenC.
DR InterPro: IPR001137; GH_11.
```

```
DR InterPro: IPR000757; Glyco_hydro.16.
DR Pfam: PF02018; CBM_4.9; 2.
DR Pfam: PF00457; Glyco_hydro.11; 2.
DR Pfam: PF00722; Glyco_hydro.16; 2.
DR PRINTS: PR00911; GLHYDRLASE11.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS00776; GLYCOSTYL_HYDROL_F11.1; 2.
DR PROSITE: PS00777; GLYCOSTYL_HYDROL_F11.2; 2.
DR PROSITE: PS01034; GLYCOSTYL_HYDROL_F16; 2.
FT CHAIN 2
FT FAMILY 11 XYLANASE /FAMILY 16 BETA
FT (1,3-1,4) GLUCANASE.
FT F36BC68805FC5274 CRC64;

SEQUENCE 802 AA: 89019 MW; 1E9D8BFA430DEBA5 CRC64;

Query Match 13.3%; Score 242.5; DB 2; Length 802;
Best Local Similarity 34.9%; Pred. No. 3.7e-08;
Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

OY 30 FSGAELTYLEVOYGFKFEARMKMAASGVSSMFLYONGSEIADGRPWVEVDIEVGKNP 89
Db 634 YSGGEFRTNNFYHYGYECSSMQAMKNDGVSSFFTYTGPS---DGNPMEIDEIILGKNT 690
OY 90 GSFQSNITTKGKGAOKTSEKHNAVSPADQAEHTYGLETPVYVNTVDGQEVKRTGQ 149
Db 691 TQVQFNYVTNGQGH--EKLYDLGFDSSEAHYTYGFDWQPNYIAWYVDGREVYRA---- 743
OY 150 VSNLTGTCG-LRFNLMSSESA-AWGOFPDESKLPLOFINWVYVK 193
Db 744 TODIPKTKIMNMPGLTVDDMLKAFN-GRTPLAHYQWVYNNK 788

RESULT 7
O93GE6 PRELIMINARY: PRT: 214 AA.
AC 093GE6:
Db 01-DEC-2001 (TREMBLrel. 19, Created)
Dr 01-DEC-2001 (TREMBLrel. 19, last sequence update)
Dr 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Endo 1-3,1-4-beta-glucanase (Fragment).
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxId=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Mellado R.P.;
RT "Novel lichenases from soil."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF254961; AAK50614.2;
DR InterPro: IPR000757; Glyco_hydro.16.
DR Pfam: PF00722; Glyco_hydro.16; 1.
DR PROSITE: PS01034; GLYCOSTYL_HYDROL_F16; UNKNOWN_1.
DR NON_TER 1
FT SEQUENCE 214 AA: 24131 MW; 1E9D8BFA430DEBA5 CRC64;

Query Match 13.1%; Score 239.5; DB 2; Length 214;
Best Local Similarity 35.5%; Pred. No. 1.1e-08;
Matches 61; Conservative 24; Mismatches 76; Indels 11; Gaps 6;

OY 23 TNVSAKDFSGAELTYLEVOYGFKFEARMKMAASGVSSMFLYONGSEIADGRPWVEVDI 82
Db 52 TSPSYNKFPCGERSVQVTGYGLYEVRMKPAKNTGIVSSFFTYTGPT--GTPMDEIDI 108
OY 83 EVLGNKPGSFQSNITTKGKGAOKTSEKHNAVSPADQAEHTYGLETPVYVNTVDGQEV 142
Db 109 EFLGKDTTKVQENYYTNGAG---NHEKLDLGFDAANAHYTFADWQPSIKWYVDGQ-L 164
OY 143 RTEGGQVSNLTGQGLRFNLMSSESA-AWGOFPDESKLPLOFINWVYVK 193
Db 165 KHTATTQITPAAGK--IMNMNMGTCVDDMLGSYGNV-PLTAHIDWVAYTK 213

RESULT 8
O4S691 PRELIMINARY: PRT: 239 AA.
ID O4S691
```

045691: 01-NOV-1996 (Tremblrel, 01, Created)  
 DT 01-NOV-1996 (Tremblrel, 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)  
 DE Endo-beta-1,3-1,4-glucanase.  
 GN BEG1.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 CX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIB 8565;  
 RA van Rensburg P., van Zyl W.H., Pretorius I.S.:  
 RT "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-  
 RT glucanase gene together with the Butyrivibrio fibrisolvens endo-beta  
 RT 1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-  
 RT glucanase gene in Saccharomyces cerevisiae.";   
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U60830; AAB05759.1; -  
 DR HSSP; P27051; IG8G.  
 DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00722; Glyco\_hydro\_16\_1.  
 DR PRINTS: PR00737; GLHIDRLASE16.  
 DR PROSITE; PS01034; GLYCOSYL-HYDROL\_F16; 1.  
 QO PROSITE 239 AA; 26969 MW; C4219760D13F878A CRC64;  
 QO SEQUENCE

[illegible]

RESULT	9
Q8RMP0	
ID	Q8RMP0
AC	Q8RMP0;
DT	01-JUN-2002 (Tremblrel, 21, Created)
DT	01-JUN-2002 (Tremblrel, 21, Last sequence update)
DT	01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE	Beta-1,3-1,4-glucanase.
OS	Bacillus subtilis.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC	Bacillaceae; Bacillus.
OX	NCBI_TaxId=1423;
RN	111
RP	SEQUENCE FROM N.A.
RA	Sum J., Li W., Gu S., Xu Z., Zhao H., Xiao J.,
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF490978; AM08358.1; -
Q8	SEQUENCE 242 AA: 27238 MW: 285FED3FE76AE69A CRC64:

Query Match	13.08;	Score 236.5;	DB 2;	Length 242;
Best Local Similarly	34.98;	Pred. No. 2e-08;		
Matches 60; Conservative	25;	Mismatches	76;	Indels 11; Gaps 6

OY 23 TNVSAKDSSGAELTLEVOYGKREARKMAAASCTVSMFLVONSEINDGRWRVEVDI 82  
| : : : | : : : | : : : | : : : |  
Db 80 TSPSYNRFDCGENSESVOTGYTGLEVMPKAKNIGIVSFEFTYGP---DGTPEWDEIDI 136  
| : : : | : : : | : : : | : : : |  
OY 83 EVLGNKPGSPQSNIITGKAGAKTSKHHAAVSPPADDAFHTTGLEMTPNRVRTVDQGEV 142  
| : : : | : : : | : : : | : : : |

[illegible]

	RESULT	10
09FDC9		
ID	Q9FDC9	PRELIMINARY; PRT; 212 AA.
AC	Q9FDC9;	
DT	01-MAR-2001 (TREMELREL_16, Created)	
DT	01-MAR-2001 (TREMELREL_16, Last sequence update)	
DT	01-DEC-2001 (TREMELREL_19, Last annotation update)	
DE	Endo-1,3-1,4-beta-glucanase (Fragmant).	
OC	Paenibacillus polymyxa (Bacillus polymyxa).	
CC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales; Paenibacillaceae; Paenibacillus.	
OX	NCBI_TaxID=1406;	
RN	[1]	
RA	SEQUENCE FROM N.A.	
RP	Yao W., Wang Y., Song W., Yang K., Su Z.;	
RT	"Gene cloning of an antifungal protein.";	
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF284449; AAC02415.1; -	
DR	HSSP; P23904; IAKR.	
DR	InterPro; IPR000757; Glyco_hydro_16.	
DR	Pfam; PF00722; Glyco_hydro_16; 1.	
DR	PRINTS; PR00737; GLHYDRLASE16.	
DR	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.	
FT	NON_TER	1
FT	NON_TER	1
SO	SEQUENCE	212 AA; 24127 MW; 7E707EF7BEF7A440 CRC64;

	Query March	12.7%	Score 232.5	DB 2	Length 212
	Best Local Similarity	34.8%	Pred. No. 3,1e-08		
	Matches 62	Conservative 26	Psmatches 73	Indels 17	Gaps 9
Qy	20	ALTTVNSAKDFGALYLTLEEVOYCKFEAKRMMAASGTVSSMFLYQNSIADGRPYVE	79		
Db	48	SLTSSAYNK-FPGGEYRSNNTYRGYLEVNMKPAKNTGTVSSFFITGCP--ANGQWME	103		
Qy	80	VDIEVLNKGPSFQSNILITGKAGAACTSEKHHAAVSPADADAFHYTGLEWTPNVRWTVDG	139		
Db	104	IDIEFLGKDTTVKQVQPNYYNTGIGGH---EKVVDLGFDDASSGHHTAFAFDQPGIKWYVDG	160		
Qy	140	QEVRTTEGGQVSNLTGCTQG-LRFNLMSSES-AAMVGGQPDSESLPLPQFLINNVKVKYKT	195		
Db	161	-VLKHT---ATTNIKRTPGQIMMNLNMGTVSDMLGPPYAGVN-PLTAEYDWN---KYT	210		

RESULT	11
007856	
ID	007856
AC	007856;
DT	01-JUL-1997 (TREMblrel. 04, Created)
DT	01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE	Beta-(1,3-1,4)-glucanase precursor.
GN	LICHENASE.
OS	Streptococcus bovis.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales
OC	Streptococcaceae; Streptococcus.
NCBI_TaxID=1315;	
XX	

RA Ekinci M., Flint H.J.;  
RT "Isolation and overexpression of a gene encoding an extracellular  
beta-(1,3-1,4)-glucanase from *Streptococcus bovis* Jb1.";  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; 292911; CAB07443.1; -.  
HSSP; P23904; IAU.



DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
 DR PRINTS: PR00737; GLHYDRLASE16.  
 DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR SIGNAL.  
 KW SIGNAL.  
 FT CHAIN 1 24 POTENTIAL.  
 FT CHAIN 25 237 BETA-(1,3-1,4)-GLUCANASE.  
 SQ SEQUENCE 237 AA; 26989 MW; 7DEF5BCE53790470 CRC64;

Query Match 12.6%; Score 230; DB 2; Length 237;  
 Best Local Similarity 33.1%; Pred. No. 5.2e-08;  
 Matches 54; Conservative 26; Mismatches 69; Indels 14; Gaps 6;

OY 30 FSGAELTYLEEVQYKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDIEVLGKNP 89  
 DB 81 YTGEMRSKEREGYGLFYNNMKPIKNPGVSSFFTYTGS--DGTKMEIDIEFLGKDT 137  
 OY 90 GSFQSNITTKGAKQKTSKHAHVPADQAHHTYGLTEPTNYVRYVTVDGQEVKTEGQ 149  
 DB 138 TKVQENYYT--SGQNHETLYNLGFDASQGFHTYGFDMQAOHITWYVDGRAVYTA---- 190  
 OY 150 VSNLTGTGQ-LRFNLW--SSESAAMVGOPDESKLPLOFINNV 189  
 DB 191 YNNIPSTPEKIMMNAWPGTHEVDSWLGATN-GRTPLYATYDWI 232

RESULT 12  
 O93U11 PRELIMINARY: PRT: 205 AA.

AC O93U11.  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Endo-1,3-1,4-beta-glucanase (Fragment).  
 OC Uncultured bacterium.  
 OS Bacteria; environmental samples.  
 OX NCBI\_TaxID=77133;  
 RP [1]  
 RN SEQUENCE FROM N.A.  
 RA Rodriguez V., Mellado R.P.;  
 RT "Novel lichenases from soil."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF254960; AAK50613.1; -  
 DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
 DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; UNKNOWN\_1.  
 DR NON\_TER 1 1  
 FT NON\_TER 205 205  
 SQ SEQUENCE 205 AA; E7CADAAPF02A2F51 CRC64;

Query Match 12.2%; Score 223; DB 2; Length 205;  
 Best Local Similarity 36.1%; Pred. No. 1.3e-07;  
 Matches 56; Conservative 21; Mismatches 68; Indels 10; Gaps 5;

OY 23 TNVSAKDFSGAELTYLEEVQYKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDI 82  
 DB 52 TSPSYNKFEDCGENRSVQYGYLYEVRMKPAKNTGIVSSFFTYTGTTE--GPMDEIDI 108  
 OY 83 EYLGNPDSFGQSNITTKGAKQKTSKHAHVPADQAHHTYGLTEPTNYVRYVTVDGQEV 142  
 DB 109 ESLGDKTKKVDQENYYTNGG--NHEKTLADIGFDANAFADWQPSIKWYVDGQ-L 164  
 OY 143 RKEGGOVSNLTGTGQLRFLNWSSEA-AWVGOPD 176  
 DB 165 KHTATTQVPAAPCK--IMMNLNMGCVDDMLGSYN 197

RESULT 13  
 O97FD3 PRELIMINARY: PRT: 246 AA.  
 AC O97FD3.  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Endo-1,3(4)-beta-glucanase family 16.  
 GN CAC2807.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1488;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=2159325; PubMed=11466286;  
 RA Neelling J., Breton G., Ometchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,  
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 bacterium Clostridium acetobutylicum";  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE007778; AAK60751.1; -  
 DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
 DR PRINTS: PR00737; GLHYDRLASE16.  
 DR PROSITE: PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR Complete proteome.  
 SQ SEQUENCE 246 AA; 27717 MW; COBEC302678D2FE1 CRC64;

Query Match 12.1%; Score 221; DB 16; Length 246;  
 Best Local Similarity 33.7%; Pred. No. 2.2e-07;  
 Matches 58; Conservative 19; Mismatches 73; Indels 22; Gaps 8;

OY 30 FSGAELTYLEEVQYKFEARKMAAAGTVSSMFLYONGSEIADGRPWVEVDIEVLGKNP 89  
 DB 91 YAGGEVRSNNRRGYGLYRSMKPAKHIGVDSFFSYTGS---DNNPMEIDIEFLGKDT 147  
 OY 90 GSFQSNITTKGAKQKTSKHAHVPADQAHHTYGLTEPTNYVRYVTVDGQEVKTE 146  
 DB 148 TEVQFNYYTNGV-----KHETLYKLGFDASQGFHTYGYIMQNTYIAMLVDGKEYRA- 200  
 OY 147 GGOVSNLTGTGQ-LRFNLWSSEA-AWVGOPDESKLPLOFINNVYVYTP 196  
 DB 201 ---TSNIPTHPGKVMNMLPGIGVDSWLGATN-GYTPVRAVYVMA--MYNP 245

RESULT 14  
 O8U8N5 PRELIMINARY: PRT: 289 AA.

AC O8U8N5.  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Endo-1,3-1,4-beta-glucanase.  
 GN EXOK OR ATU4055 OR AGR-L\_1600.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.F., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Glendening J., Deatherage G., Gillet W., Grant C.,  
 RA Kutavain T., Levy J., Li W.-J., McClelland E., Palmeri A.,  
 RA Raymond S., Rouse G., Saenphitmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 C58.";  
 RL Science 294:2317-2323(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE-21608551: PubMed-11743194:  
RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,  
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Hounzel K., Gordon J., Vaudin M., Iatichouk O., Epp A., Liu F.,  
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,  
RA Cleio C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328(2001).  
DR EMBL: AE009336: AAL44856.1: ALT\_INIT.  
DR EMBL: AE008279: AAK89373.1: -  
KM Complete proteome.  
SQ SEQUENCE 289 AA: 32368 MW: 0384C6F83320EAC9 CRC64:

Query Match 11.3%; Score 206; DB 16; Length 289;  
Best Local Similarity 35.1%; Pred. No. 2.7e-06;  
Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

QY 28 KPFSCAELTYLLEVOYCKFEARKMAAASGVSMFLYONGSEIADGRPWVEVDIEVLK 87  
DB 115 RNPAGEIOTKGRYRTGYEARKKATGSLNSAFFTYIGPT--DKKPHDEIDFEVLK 171  
QY 88 NPGSFQSN-IITGKAGAKTSEKHHAVSPADQAFHTYGLEWTPNYVRMTVDGOEVRKTE 146  
DB 172 NTGKQVNLNGYIAKGG---NEKLVPYEGADAGFNDYAFVMEPRRLRYVNGKLVH--- 224  
QY 147 GCGVSNLTG---GTGGLRPNLWSSESA-AWVGOF 175  
DB 225 --EVTDETKIPQNAOKIFFSLMGCTDLKDMGKF 256

## RESULT 15

Q98C78 PRELIMINARY: PRT; 293 AA.  
AC Q98C78;  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE Endo-1,3-1,4-beta-glycanase, EXOK.  
GN MLR5264.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OK NCBI\_Taxid=381:  
RX [1]  
KM SEQUENCE FROM N.A.  
RP STRAIN-MAFF303099;  
RC MEDLINE-21082930: PubMed-11214968;  
RX Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP003006: BAB51743.1: -  
DR InterPro: IPR00757; Glyco\_hydro\_16.  
DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
DR PRINTS: PR00737; GLHYDRASE16.  
KM Complete proteome.  
SQ SEQUENCE 293 AA: 32347 MW: 938BA1672176ED3C CRC64:

Query Match 11.0%; Score 200; DB 16; Length 293;  
Best Local Similarity 33.8%; Pred. No. 7e-06;  
Matches 51; Conservative 27; Mismatches 61; Indels 12; Gaps 6;

QY 28 KPFSCAELTYLLEVOYCKFEARKMAAASGVSMFLYONGSEIADGRPWVEVDIEVLK 87  
DB 112 REFACGEIOTKORFGYGYEARLKTDTGSLNAAFTYIGPS---DKQPMDEIDFEILTK 168

QY 88 NPGSFQSN-IITGKAGAKTSEKHHAVSPADQAFHTYGLEWTPNYVRMTVDGOEVRK-T 145  
DB 169 DTSKQVNAVYIOCKGKNEKLEVE---VPGCTDKAFNDYAFVMEKDSLRYVNGQLVNTIT 224  
QY 146 ECGVSNLTGTGGLRPNLWSSSES-AAWVGOF 175  
DB 225 DPAKLP--SHAQKIFFSLMGSEIMKGWMAF 253

Search completed: January 9, 2003, 12:18:13  
Job time : 32.103 secs

DR	PDB; 1CPN; 22-JUN-94.	24	237	BETA-GLUCANASE.
DR	PDB; IMAC; 27-FEB-95.	128	128	NUCLEOPHILE (BY SIMILARITY).
DR	PDB; 1AJK; 06-MAY-98.	132	132	PROTON DONOR (BY SIMILARITY).
DR	PDB; 1AJO; 06-MAY-98.	55	84	
DR	InterPro; IPR000757; Glyco_hydro_16.	128	128	
DR	Pfam; PF00722; Glyco_hydro_16; 1.	26	26	
DR	PRINTS; PR00737; GLHYDRLASE16.	29	31	
DR	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.	41	43	
KW	Hydrolase; Glycosidase; Signal; 3D-structure.	51	52	
KW	SIGNAL	53	53	
FT	CHAIN	56	57	
FT	ACT_SITE	59	61	
FT	ACT_SITE	62	64	
FT	DISULFID	66	67	
FT	MUTAGEN	70	78	
FT	STRAND	79	80	
FT	STRAND	81	89	
FT	STRAND	93	93	
FT	STRAND	96	103	
FT	TURN	108	109	
FT	STRAND	110	118	
FT	HELIX	120	122	
FT	TURN	123	123	
FT	STRAND	127	134	
FT	TURN	135	136	
FT	STRAND	137	147	
FT	TURN	148	149	
FT	STRAND	150	150	
FT	STRAND	155	158	
FT	TURN	163	164	
FT	STRAND	168	174	
FT	STRAND	179	183	
FT	TURN	184	185	
FT	STRAND	186	191	
FT	STRAND	200	209	
FT	TURN	213	216	
FT	STRAND	225	236	
SO	SEQUENCE	237	AA: 26589 MW: 436EABCDFFC87781 CRC64:	

```
Query Match      .    12.5%; Score 228; DB 1; Length 237;
Best Local Similarity 34.3%; Pred. No. 1.5e-09;
Matches          60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;
```

QY	23	TNYSAKDFSCAELYTLEEVOYGKFEARKMAAAAGCTVSSMFLYONGSIADGRPWVEVDI	82
		: : :       :   :   :   :           :     :	
Dd	75	TSSAYNKRFDCAEYNSTNIYGYGLYEVMKPAKNITGVSSFSTYTGPK---AHGTOWDEIDI	131
		:     :   :   :                 :	
QY	83	EVLCKNGPGSGNIITCKAGAOKTSEKHHAVSAPDAQAFTHTGLEWTPNYRVRTWDGOEV	142
		:     :   :   :                 :	
Dd	132	EFLGKDITTKVFNYNTNGVGGH---EKVISLGFDPASKGFTHTYAEDQWGKYIKWYDVG-VL	187
		:     :   :   :                 :	
QY	143	RKTGGGVSNLTGTQG-LRFLNLWSES-AWVGQFDESKLPFLFOFINVKVKKYT	195
		:   :   :       :   :   :   :   :   :   :	
Dd	188	KHT---ATANIPSPGKIMLNNGTGDVDMWLSYNGAN-PLYAEYDW---KVT235	
		:   :       :   :   :   :   :   :   :	

RESULT 10			
ID	GUB_PAEPO	STANDARD;	PRT; 238 AA.
AC	P45797;		
DT	01-NOV-1995	(Rel. 32, Created)	
DT	01-NOV-1995	(Rel. 32, Last sequence update)	
DT	15-JUL-1998	(Rel. 36, Last annotation update)	

DE	Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE	(1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN	GLUB.
OS	Paenibacillus polymyxa (Bacillus polymyxa).
OC	Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
OX	NCBI_TaxId=1406;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 842;
RC	MEDLINE=92041687; PubMed=1938968;
RA	Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
RR	"Two beta-glucanase genes are clustered in Bacillus polymyxa:
RT	molecular cloning, expression, and sequence analysis of genes
RT	encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
RT	J. Bacteriol. 173:7705-7710(1991).
RL	-!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC	in beta-D-glucans containing 1,3- and 1,4-bonds.
CC	-!- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC	SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC	-!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).

DR	EMBL; X57094; CAA40379.1; -
DR	HSSP; P23904; ICPN.
DR	InterPro; IPR000757; Glyco_hydro_16.
DR	Pfam; PF00722; Glyco_hydro_16; 1.
DR	PRINTS; PR00737; GLHYDRLASE16.
DR	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW	Hydrolase; Glycosidase; Signal.
	1 26
FT	SIGNAL
FT	CHAIN
FT	27 238
FT	BETA-GLUCANASE.
FT	ACT_SITE 129 129
FT	NUCLEOPHILE (BY SIMILARITY).
FT	ACT_SITE 133 133
FT	PROTON DONOR (BY SIMILARITY).
FT	DISULFID 56 85
FT	BY SIMILARITY.
SQ	SEQUENCE 238 AA; 26919 MW; C0CF7B4EA5D40B8C CRC64;

	Query Match	12.5%; Score 227.5; DB 1;	Length 238;
	Best Local Similarity	34.3%; Pred. No. 1.6e-09;	Matches
	Matches	61; Conservative	26; Mismatches 74; Indels 17; Gaps 9;
Qy	20 ALTNYSAKDFSGAELYLTLEEVGYKFEARKMWAASGTVSSMFELXONGSEIADGRPWE	79 :    :	:
Dd	74 SLTSPANNK-FDCGEYRSTNNYGGLVEVSMKPAAKTNGIVSFFTYTGPSH--GTQMDE	139 :    :	:
Qy	80 VDEVLCNKGCSFQSNIITCKAQAGTKSEKHAVSPAADAQAFTYLGLEWTPNVVRVTVDG	139 :       :	:
Dd	130 IDIEFLGGDKTTKVQNYYTNCVGSH---EKLINLGFDASTSFHTYAEDWDQPGVIKKYVDG	186 :       :	:
Qy	140 QEVKRTEGGGVSNLTIQTQG-LRFNLWSESA-AWWGFODESKLPFLFINMKVKYKYT	195 :    :	:
b	187 -VLKHT---ATNTNPSPDGKITMNWNIGCTGSDNSIKSYNGCAN-PLYAFYPDVV---KYT	236 :    :	:

RESULT	11
EXOK_RHIME	
ID	EXOK_M
AC	P33693
DT	01-FE
DT	01-FE
DT	15-JUN
DE	Endo-1
DE	biosyn
GN	EXOK
OS	Rhizo
OG	Plasm
OC	Bacter

Query Match 10.8%; Score 197; DB 1; Length 269;  
Best Local Similarity 32.5%; Pred. No. 2.8e-07;  
Matches 50; Conservative 31; Mismatches 61; Indels 12; Gaps 6;



```

FT CONFLICT 234 234 S -> C (IN REF. 3).
SQ SEQUENCE 347 AA; 36092 MW; 1E8AFB862C4BB328 CRC64;

Query Match
Best Local Similarity 10.4%; Score 189.5; DB 1; Length 347;
Matches 83; Conservative 49; Mismatches 120; Indels 85; Gaps 15;

QY 28 KDFSGAELTYLEVOYGRFEARMKMAASGVSSMFLYQNGSEIADGRPWVEVDIEVLGK 87
   I : : : : : I I I I I : : : : : I I I I I : : : : :
Dh 3 KNSGCTVLSSTRAVWYGVKSARIKTSHLAGVVTGILYSGAGD-----ELDYEFVGA 54
   I : : : : : I I I I I : : : : : I I I I I : : : : :

QY 88 NPGSFQSN-----ITGKAGAQTSEKHAVSPAQAQAFHTYGLEWTPNPNVVRVTDGQ 140
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
Dh 55 DLETAQTFYVESVLYNYSANISTD-----TFENVHTYELDWHEDYVVTWSIDGV 105
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :

QY 141 EVRKTEGQVSNLTCTOGLRF-----NLW-----SSESA-----AWVG-----QDESCLP 181
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
Dh 106 VGRTRYKKETYNAT-TQYQVPTQPSKVDISIWPGGNSINAPGTTAWSGEINWDASDIS 164
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :

QY 182 ----LQFINWKKYKTPGCGEGSGDFTLDWTNFDFTDGRWGKGWDFTDGNRVDLTD 237
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
Dh 165 NPGYIYAIVNEVNIICYDP-----PSDTKKNGTSAVYVTSSE-----FLAKDIAITD 212
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :

QY 238 KNIY--SRDGMILALTKRGESFNGQVRDDEPAPQSSSSAPASSSVPAS----- 287
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
Dh 213 DEVMDSDG-----SGLDPHKGTAT-----TSSTQKSSSSSTATSSKTSDDHSSSTKK 260
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :

QY 288 ---SSSVPASSSAFVPSSSSSATNAIHGMRTTPAVA 321
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
Dh 261 SKTSTASSSSSSSSSSSSSTATKNGDKVSVSVS 297
   : : : : : I : : : : : I : : : : : I : : : : : I : : : : :

RESULT 13
MER5_ARATH
ID MER5_ARATH STANDARD; PRT; 269 AA.
AC P24806; Q39148; Q41904; O64956;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE MERI-5 protein precursor (Endo-xyloglucan transferase) (Xyloglucan
endo-1,4-beta-D-glucanase)
GN MERI-5 OR MERI5B OR SEM4 OR AT4G30270 OR F9N11.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_taxid=3702;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93005704; PubMed=1840916;
RA Medford J.I., Elmer J.S., Klee H.J.;
RT "Molecular cloning and characterization of genes expressed in shoot
apical meristems."
RL Plant Cell 3:359-370(1991).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,

```

```

RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Gymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Draefoer E.,
RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braum M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Koster W.,
RA Moolijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Blocker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefod F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana";
RL Nature 402:769-777(1999).
[5]
RN SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RT-PCR Arabidopsis full length cDNA clones (RAFTs) sequenced by the
SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE OF 1-132 FROM N.A.
RX STRAIN=cv. Columbia; TISSUE=Green siliques;
RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
[7]
RN SEQUENCE OF 1-120 FROM N.A.
RX STRAIN=cv. Columbia; TISSUE=Leaf;
RA MEDLINE=98278374; PubMed=9617812;
RA Park J.-H., Oh S.A., Kim Y.H., Woo H.R., Nam H.G.;
RT "Differential expression of senescence-associated mRNAs during leaf
senescence induced by different senescence-inducing factors in
Arabidopsis."
RL Plant Mol. Biol. 37:445-454(1998).
CC -I- FUNCTION: Involved in cell wall reconstruction.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN SHOOT APICAL MERISTEMS, ALSO
FOUND IN SEEDLINGS AND MERISTEMS.
CC -I- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -I- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
in positions 158; 178; 183; 189; 190; 194 and 199.
CC -I- CAUTION: Ref.6 sequence differs from that shown due to frameshifts
in positions 93 and 104.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

222

Query Match	7.0%;	Score 128;	DB 1;	Length 283;
Best Local Similarity	23.7%;	Pred. No. 0.026;		
Matches 46;	Conservative 41;	Mismatches 73;	Indels 34;	Gaps 9;

Search completed: January 9, 2003, 12:16:49  
Job time : 9.88657 secs

1

1



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 11:50:42 ; Search time 115.824 Seconds  
(without alignments)  
1368.672 Million cell updates/sec

Title: US-09-654-652A-1  
Perfect score: 1333  
Sequence: 1 MVSADFGAELTYLEEVOY.....TRKQESFNGVPRDDEPAP 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				Pending_Patents_AA_Main:*	
1:	/cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*	2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*	3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US080_COMB.pep.*	5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*	6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*	8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*	9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*	11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*	12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*	14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*	15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*	17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*	18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*	20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*	21:	/cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*	23:	/cgn2_6/ptodata/1/paa/US099_COMB.pep.*	24:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*	26:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*	27:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	1333	100.0	248	20	US-09-654-652A-1
2	1333	100.0	267	20	US-09-654-652A-2
3	1328	99.6	349	20	US-09-654-652A-3
4	1328	99.6	349	21	US-09-791-537-30441
5	272.5	20.4	214	21	US-09-791-537-145019
6	266.5	20.0	259	21	US-09-791-537-106873

7	265.5	19.9	252	21	US-09-791-537-32335	Sequence 32335, A
8	257.5	19.3	308	18	US-09-463-862A-1	Sequence 1, Appli
9	256.5	19.2	208	21	US-09-791-537-6910	Sequence 6910, Ap
10	250.5	18.8	394	21	US-09-791-537-132244	Sequence 132244,
11	247.5	18.6	214	21	US-09-791-537-117361	Sequence 117361,
12	247.5	18.6	276	21	US-09-791-537-31771	Sequence 31771, A
13	246.5	18.5	242	21	US-09-791-537-27063	Sequence 27063, A
14	245.5	18.4	242	21	US-09-791-537-73450	Sequence 73450, A
15	245.5	18.4	242	21	US-09-791-537-79889	Sequence 79889, A
16	245	18.4	239	21	US-09-791-537-103347	Sequence 103347,
17	244	18.3	239	21	US-09-791-537-103324	Sequence 103324,
18	242.5	18.2	802	21	US-09-791-537-50107	Sequence 50107, A
19	242.5	18.2	802	21	US-09-791-537-116925	Sequence 116925,
20	238.5	17.9	239	21	US-09-791-537-30436	Sequence 30436, A
21	237	17.8	334	21	US-09-791-537-48909	Sequence 48909, A
22	235.5	17.7	214	21	US-09-791-537-88552	Sequence 88552, A
23	234.5	17.6	239	21	US-09-791-537-18679	Sequence 18679, A
24	233	17.5	239	21	US-09-791-537-103817	Sequence 103817,
25	231.5	17.4	320	21	US-09-791-537-72048	Sequence 72048, A
26	230	17.3	237	21	US-09-791-537-72848	Sequence 72848, A
27	229.5	17.2	243	21	US-09-791-537-30440	Sequence 30440, A
28	226	17.0	212	21	US-09-791-537-88873	Sequence 88873, A
29	226	17.0	214	21	US-09-791-537-76270	Sequence 76270, A
30	226	17.0	237	21	US-09-791-537-103320	Sequence 103320,
31	226	17.0	237	21	US-09-791-537-106876	Sequence 106876,
32	226	17.0	239	21	US-09-791-537-103345	Sequence 103345,
33	226	17.0	239	21	US-09-791-537-103350	Sequence 103350,
34	226	17.0	239	21	US-09-791-537-103355	Sequence 103355,
35	226	17.0	239	21	US-09-791-537-103357	Sequence 103357,
36	226	17.0	239	21	US-09-791-537-103814	Sequence 103814,
37	226	17.0	239	21	US-09-791-537-103819	Sequence 103819,
38	226	17.0	240	21	US-09-791-537-64827	Sequence 64827, A
39	225	16.9	238	21	US-09-791-537-79887	Sequence 79887, A
40	222.5	16.7	242	21	US-09-791-537-37095	Sequence 37095, A
41	222	16.7	237	21	US-09-791-537-3485	Sequence 3485, Ap
42	214	16.1	160	21	US-09-706-729-14	Sequence 14, Appl
43	206	15.5	220	21	US-09-739-449-8386	Sequence 8386, Ap
44	206	15.5	220	22	US-09-803-110-8386	Sequence 8386, Ap
45	204.5	15.3	504	27	US-60-385-568-329	Sequence 329, App

ALIGNMENTS

RESULT 1  
US-09-654-652A-1  
; Sequence 1, Application US/09654652A  
; GENERAL INFORMATION:  
; APPLICANT: SHYUR, LIE-FEN  
; APPLICANT: CHEN, JUI-LIN  
; APPLICANT: YANG, NING-SUN  
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBLAST SUCCINOGENES 1, 3-1,  
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMIC ACTIVITY AND  
; TITLE OF INVENTION: THERMO-TOLERANCE  
; FILE REFERENCE: 4910-8  
; CURRENT APPLICATION NUMBER: US/09/654,652A  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 248  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme  
; OTHER INFORMATION: with enhanced activity and thermal stability  
US-09-654-652A-1

Query Match 100.0%; Score 1333; DB 20; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3.4e-130;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 MVSADFGAELTYLEEVOYCKFEARMKMAASGTSSMFLYQNGSEIADGRPWVEVDIE 60

```

|||||
Db 1 MVSADFGSAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60
QY 61 VLGNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 120
Db 61 VLGNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 120
QY 121 KTEGQVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 180
Db 121 KTEGQVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 180
QY 181 DFTLWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 240
Db 181 DFTLWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 240
QY 241 PRDDEPAP 248
Db 241 PRDDEPAP 248

RESULT 2
US-09-654-652A-2
; Sequence 2, Application US/09654652A
; GENERAL INFORMATION:
; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1.
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMIC ACTIVITY AND
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654.652A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified enzyme
; OTHER INFORMATION: with enhanced activity and thermal stability
US-09-654-652A-2

Query Match 100.0%; Score 1333; DB 20; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.8e-130;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVSADFGSAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60
Db 1 MVSADFGSAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 60
QY 61 VLGNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 120
Db 61 VLGNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 120
QY 121 KTEGQVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 180
Db 121 KTEGQVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 180
QY 181 DFTLWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 240
Db 181 DFTLWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 240
QY 241 PRDDEPAP 248
Db 241 PRDDEPAP 248

RESULT 3
US-09-654-652A-3
; Sequence 3, Application US/09654652A
; GENERAL INFORMATION:

```

```

; APPLICANT: SHYUR, LIE-FEN
; APPLICANT: CHEN, JUI-LIN
; APPLICANT: YANG, NING-SUN
; TITLE OF INVENTION: A TRUNCATED FORM OF FIBROBACTER SUCCINOGENES 1, 3-1.
; TITLE OF INVENTION: 4-BETA-D-GLUCANASE WITH IMPROVED ENZYMIC ACTIVITY AND
; TITLE OF INVENTION: THERMO-TOLERANCE
; FILE REFERENCE: 4910-8
; CURRENT APPLICATION NUMBER: US/09/654.652A
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 349
; TYPE: PRP
; ORGANISM: Fibrobacter succinogenes
US-09-654-652A-3

Query Match 99.6%; Score 1328; DB 20; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.9e-129;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSAKDFSGAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 61
Db 25 VSAKDFSGAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 84
QY 62 LGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 121
Db 85 LGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 144
QY 122 TEGGOVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 181
Db 145 TEGGOVSNLTGTQGLRNLWSSSAWVGQFDESKLPFQFINNVKVKYKTPGQEGGS 204
QY 182 FTLDWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 241
Db 205 FTLDWNTDNFTFDGSRGKGDWTFDGNRVLDLTKNIYSRDGMLILALTRKQSFNGQV 264
QY 242 RDDEPAP 248
Db 265 RDDEPAP 271

RESULT 4
US-09-791-537-30441
; Sequence 30441, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30441
; LENGTH: 349
; TYPE: PRP
; ORGANISM: Fibrobacter succinogenes
US-09-791-537-30441

Query Match 99.6%; Score 1328; DB 21; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.9e-129;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSAKDFSGAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 61
Db 25 VSAKDFSGAELYTLEEVQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIE 84
QY 62 LGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 121
Db 85 LGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFTTYGLEWTPNYVRWTVDGQEV 144

```



```
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Bacillus alkalophilus DSM 9956
US-09-463-862A-1

Query Match      19.3%; Score 257.5; DB 18; Length 308;
Best Local Similarity 36.5%; Pred. No. 1.2e-17;
Matches 62; Conservative 22; Mismatches 73; Indels 13; Gaps 6;

QY 6 DFGSGLYTLLEEVQCKFEARKMAAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
Db 96 EYKAGELRNQYQYGLFEVNMKPAKSTGTVSSLFY-TGPDWDNDPWEIDIEFLGKD 154
QY 66 PGFSQNIITGKAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGG 125
Db 155 TTRVQFNFTNGVG---NNEHYHELGFDASEFNYYAFENRPEISWYNGVELVYTA--- 208
QY 126 QVSNLTGT-QGLRNLWSSESA-AWVGQFDESKLPLFQFINNVKVKYKTP 173
Db 209 -TENIPQTPQKIMMNLWPGIGVDGWTGVPDGEDTPVVTEDWV---RYTP 254

RESULT 9
US-09-791-537-6910
; Sequence 6910, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6910
; LENGTH: 208
; TYPE: PRT
; ORGANISM: pdb 1CPN
US-09-791-537-6910

Query Match      19.2%; Score 256.5; DB 21; Length 208;
Best Local Similarity 33.8%; Pred. No. 8.8e-18;
Matches 77; Conservative 29; Mismatches 95; Indels 27; Gaps 13;

QY 7 FSCAELYTLLEEVQCKFEARKMAAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db 1 FDCAYRSTNIYGYGLYEVSMPKPAKNTGIVSFFTYTGP---AHGTQWDEIDIEFLGKDT 57
QY 67 GSFSQNIITGKAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGGQ 126
Db 58 TKVQFNFTNGVG---EKVISLGFDAKSGFHTYAFDQPGYIKWYVDG-VLKHT---A 110
QY 127 VSNLTGTG-LRFNLWSSESA-AWVGQFDESKLPLFQFINNVKVKYKTP 184
Db 111 TANIPSTPGKIMMNLWNGTGVDDMLGSYNGAN-PLYAEYDWYKTSNASTDYQWNTDGG 169
QY 185 DWTDFNEDFDGSRWKGDTWFDGCRVDLT--DKNI-YSRDGMILIALT 229
Db 161 -WEPR-SYFNPSTWEKADGYSNGVFNCTWRANNVNFNTDGLKLGLT 206

RESULT 10
US-09-791-537-132244
; Sequence 132244, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
```

```
; CURRENT APPLICATION NUMBER: US/09791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132244
; LENGTH: 394
; TYPE: PRT
; ORGANISM: pdb 1AJKA
US-09-791-537-132244

Query Match      18.8%; Score 250.5; DB 21; Length 394;
Best Local Similarity 29.6%; Pred. No. 9.3e-17;
Matches 68; Conservative 33; Mismatches 90; Indels 39; Gaps 10;

QY 7 FSGAELYTLLEEVQCKFEARKMAAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db 1 FDCAYRSTNIYGYGLYEVSMPKPAKNTGIVSFFTYTGP---AHGTQWDEIDIEFLGKDT 57
QY 67 GSFSQNIITGKAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGGQ 126
Db 58 TKVQFNFTNGVG---EKVISLGFDAKSGFHTYAFDQPGYIKWYVDG-VLKHT---A 110
QY 127 VSNLTGTG-LRFNLWSSESA-AWVGQFDESKLPLFQFINNVKVKYKTP 175
Db 111 TANIPSTPGKIMMNLWNGTGVDDMLGSYNGAN-PLYAEYDWYKTSNASTDYQWNTDGG 169
QY 176 G-----EGSDFTLDWTDNFDTPDGSRWKGD-----WTFDGN 208
Db 170 GIVNAVSGGNYSVNWSNTGTVFVVGKGTTCSPRTTINYNAGVWAPNGN 219

RESULT 11
US-09-791-537-117361
; Sequence 117361, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 117361
; LENGTH: 214
; TYPE: PRT
; ORGANISM: pdb 1AJKA
US-09-791-537-117361

Query Match      18.6%; Score 247.5; DB 21; Length 214;
Best Local Similarity 32.9%; Pred. No. 8e-17;
Matches 69; Conservative 32; Mismatches 84; Indels 25; Gaps 12;

QY 33 SGTSSMFLYQNGSEIADGRPWVEVDIEVLGKNPQSFOSNIITGKAGAKTSEKHHAVSP 92
Db 2 TGISSFFTYTGP---AHGTQWDEIDIEFLGKDTTKVQFNFTNGVGGH---EKVISLGF 55
QY 93 AADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGGQVSNLTGTG-LRFNLWSSESA-AWVG 150
Db 56 DASKGFHTYAFDQPGYIKWYVDG-VLKHT---ATANIPSTPGKIMMNLWNGTGVDDWLG 111
QY 151 QFDESKLPLFQFINNVKVKYKTPGQEGGSDFTLDWTDNFDTPDGSRWKKGDTWFDGCRV 210
Db 112 SYNGAN-PLYAEYDWV---KYTSNQ-TGGSFF-----EPFNSYNSGTWEKADGYSNGGVF 161
QY 211 DLT--DKNI-YSRDGMILIALTRKGOESFN 237
Db 162 NCTWRANNVNFNTDGLKLGLTSSAYNKFD 191

RESULT 12
```

```
FILE REFERENCE: 5914-090-999
CURRENT APPLICATION NUMBER: US/10/042.417
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 60/260,179
PRIOR FILING DATE: 2001-01-5
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 422
TYPE: PRT
ORGANISM: Homo sapiens
US-10-042-417-4
```

```
Query Match
Best Local Similarity 22.1%; Score 83.5; DB 12; Length 422;
Matches 30; Conservative 21; Mismatches 48; Indels 37; Gaps 4;
```

```
Qy 103 LEWTPNVYRWTVDGGVNRKTEGGVSNL-----TGVOGLRFLNLSSE 144
Db 117 LHMKVYIKAILRMKQLBDEHAFETSSLIGSHARVYALYKDLCTGSDLSAKLMQVS 176
Qy 145 SAAMV-----GQFDESKLPLFOFINWYKYKTPG-----QEGGSDFTLDMTD 188
Db 177 TGQCVYGIQHTCAVAKDEQKLVTSFQNTVACWESSGARTQHRGHTGAVFSVDYND 236
Qy 189 NEDFDGSRMGKQDWT 204
Db 237 ELDLIVS---GSADPT 249
```

```
RESULT 13
US-09-841-132-192
Sequence 192, Application US/09841132
Patent No. US20020061848A1
```

```
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841.132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 192
LENGTH: 848
TYPE: PRT
ORGANISM: Chlamydia
US-09-841-132-192
```

```
Query Match
Best Local Similarity 20.4%; Score 83.5; DB 10; Length 848;
Matches 50; Conservative 34; Mismatches 82; Indels 79; Gaps 12;
```

```
Qy 8 SGAEIYLTLEVOYG-----KFEARMKMAASGTVSSMFL-----YONGSEIA 49
Db 599 AGVDIQLMEDEFVLGVSAGAFGLKMDQKFDPAEVSRRKGVGSYTGFGLGSMFFKQYSLG 658
Qy 50 DCRPVEVDIEVLGNKPSFQSNITTKAGAQKTSKHHAVSPAADOAFHTYGLEMTNY 109
Db 659 ETQNDMKTRYGVLGESSASWTSRGLV---ADALVEYRSLVGPVPRPT---FYALHFN-P-Y 710
Qy 110 VR-----WTVDGGVNRKTEGGVSNLGTGGLRFLNLSSESAAMVGOFPDESKLPL 159
Db 711 VEVSYSKMKPFQFTGQGEARSEFEDASLTNITITPLGKMFEL---AFIKQGFSE----- 760
Qy 160 FOFIN-----WVKVYKTPGQG---EGGSDF-----TLDMTQNF 190
Db 761 ---VNSLGISYAW-EAYKRVGEGAVQLLEAGFDWEGADMDLPRQELRYALENNFTWSSYF 816
Qy 191 DTFDG 195
```

```
Db 817 STVLG 821
```

```
RESULT 14
US-09-841-132-178
Sequence 178, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
```

```
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841.132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 178
LENGTH: 1530
TYPE: PRT
ORGANISM: Chlamydia
US-09-841-132-178
```

```
Query Match
Best Local Similarity 20.4%; Score 83.5; DB 10; Length 1530;
Matches 50; Conservative 34; Mismatches 82; Indels 79; Gaps 12;
```

```
Qy 8 SGAEIYLTLEVOYG-----KFEARMKMAASGTVSSMFL-----YONGSEIA 49
Db 1281 AGVDIQLMEDEFVLGVSAGAFGLKMDQKFDPAEVSRRKGVGSYTGFGLGSMFFKQYSLG 1340
Qy 50 DCRPVEVDIEVLGNKPSFQSNITTKAGAQKTSKHHAVSPAADOAFHTYGLEMTNY 109
Db 1341 ETQNDMKTRYGVLGESSASWTSRGLV---ADALVEYRSLVGPVPRPT---FYALHFN-P-Y 1392
Qy 110 VR-----WTVDGGVNRKTEGGVSNLGTGGLRFLNLSSESAAMVGOFPDESKLPL 159
Db 1393 VEVSYSKMKPFQFTGQGEARSEFEDASLTNITITPLGKMFEL---AFIKQGFSE----- 1442
Qy 160 FOFIN-----WVKVYKTPGQG---EGGSDF-----TLDMTQNF 190
Db 1443 ---VNSLGISYAW-EAYKRVGEGAVQLLEAGFDWEGADMDLPRQELRYALENNFTWSSYF 1498
Qy 191 DTFDG 195
Db 1499 STVLG 1503
```

```
RESULT 15
US-10-007-693-98
Sequence 98, Application US/10007693
Patent No. US20020146776A1
GENERAL INFORMATION:
```

```
APPLICANT: Bhatia, Ajay
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.515C2
CURRENT APPLICATION NUMBER: US/10/007.693
CURRENT FILING DATE: 2001-12-05
NUMBER OF SEQ ID NOS: 157
SEQ ID NO 98
LENGTH: 1531
TYPE: PRT
ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-98
```

```
Query Match
Best Local Similarity 20.4%; Score 83.5; DB 12; Length 1531;
Matches 50; Conservative 34; Mismatches 82; Indels 79; Gaps 12;
```

```
Qy 8 SGAEIYLTLEVOYG-----KFEARMKMAASGTVSSMFL-----YONGSEIA 49
```

DB 292 DTGASYISG-----STSIKLEMEAL 312

RESULT 10  
US-09-738-363-4  
; Sequence 4, Application US/09738363  
; Patent No. US20010010932A1  
; GENERAL INFORMATION:  
; APPLICANT: Schuepf, Harry E.  
; Payne, Jewel M.  
; Marva, Kenneth E.  
; Foncerrada, Luis  
; TITLE OF INVENTION: Nematocidal Proteins  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jay M. Sanders  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/09/738,363  
; APPLICATION NUMBER: US/09/738,363  
; FILING DATE: 15-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/076,137  
; FILING DATE: 12-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay  
; REGISTRATION NUMBER: 39,355  
; REFERENCE/DOCKET NUMBER: MA-20CCCD3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1289 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: YES  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: BACILLUS THURINGIENSIS  
; INDIVIDUAL ISOLATE: PS17  
; IMMEDIATE SOURCE:  
; CLONE: E. coli NM522(PMYC 1628) NRRL B-18652  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-738-363-4

Query Match 5.9%; Score 85.5; DB 10; Length 1289;  
Best Local Similarity 23.2%; Pred. No. 11;  
Matches 56; Conservative 31; Mismatches 89; Indels 65; Gaps 13;

QY 33 SGTGVSMTFLYONGSEIA-----DGRPWVEVD-IEVIGKMPGSGFOSNIITGKAGAKTSTK 86  
DB 1080 NGSIRSDISYQNTDAVLPTLPKLRHWMSDRFSEGDIMAFQGLNRAVYQLEQNTLL 1139  
QY 87 HHAVSPADQAQHTYGLWTVPNVVRYTVDGQEVARKTEGGVSNLTGGLRPNLWSS--E 144  
DB 1140 HNG-----HTTKDAANKWTG-----DAHQVYLEDGKRYLRPLDMSSSVS 1179  
QY 145 SAAMVGQDESKLPLFQFINWVKYKYPGQEGGSDFTLD-----WTDNFDFPD 194

DB 1180 QTIENFDPDK--EYQLV-----FHQGEQ--TYVLEHGEETKYLETTHHFANFT 1227  
QY 195 GSRMGKDWTFPGNRYDLTDKNLYSRDGLI--LALTRKGGESFNGOVRDEPAPNSS 251  
DB 1228 TSO--RQGLTFESKNYTVI---ISSEDEGLVLDNALV-----EAPLPDQNSGNT 1275  
QY 252 S 252  
DB 1276 A 1276

RESULT 11  
US-09-738-626-6377  
; Sequence 6377, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 6377  
; LENGTH: 714  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6377

Query Match 5.8%; Score 84; DB 9; Length 714;  
Best Local Similarity 22.5%; Pred. No. 6.9;  
Matches 51; Conservative 36; Mismatches 102; Indels 38; Gaps 11;

QY 7 FSGAELTYLLEVOYGFKAEMKMAAS-----GTVSMFLYONGSEIA--DGRPWVEVDI 59  
DB 436 YSSARVH-LPEIPACNFRLTGVARARQSEELVDGVAPRAIMQNTNFCADNDGRRPGELDI 494  
QY 60 -EVLGNKPGSFOSNIITGKAG--AOKTSEKHNAVSPAADAQHTYGLWTVPNVRYRTVD 115  
DB 495 TFEYSSRVNTQYSAVHLGCAGNRPKLRQMEMESMGCD--WHDVGVEVFDGQIVFTID 552  
QY 116 GOEV-----RKTEGGVSNLTG--TQGLRFMLWSSSAAMVGQDESKLPLFQFINWVKYK 170  
DB 553 GKAIVTSSGKDVGVNSVTPPAADLRPAHFRKLESEYREVYGO-----FWHLITNTM----- 602  
QY 171 YPGQEGGSDFTLDWNTDNFTFDGSRMGKDWTFDGNRVDTDKNI 217  
DB 603 ----VQSGKDSWITAVDNNEAFPEHR-----FQIDHVAVIDESDSV 640

RESULT 12  
US-10-042-417-4  
; Sequence 4, Application US/10042417  
; Patent No. US20020123082A1  
; GENERAL INFORMATION:  
; APPLICANT: Pagano, M.  
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
; PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS



```

Db 168 QEDWEGHODDIDYDMLNLAHVKEKNGGVKRPKMYPEOLNKRMAPDSKD--FHITG 225
Oy 103 LEWTPYVRYMTVDGQEVKRTGEG---GOVSHLTGTGRLR---FNLMSSESAAMVGOFDE-- 154
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 CEVNOHEIITMYUDGVAVKARKPKYKMRPMNVITSLGRLRFYVGFDDKNNAIIPETDARA 285
Oy 155 ---SKLPLGFQFINWVKYKYPG 174
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 REKLSDIPTSMYVDYVRVWEKSAG 309

RESULT 4
US-09-731-221-79
; Sequence 79, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Atrachis
US-09-731-221-79

Query Match
Best Local Similarity 54.1%; Pred. No. 0.14; Length 526;
Matches 20; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

Oy 231 KGSEFNGQVPRDDPPAPNSSVDKLAALLENHHNNH 267
Db 491 KNNNPFKFFVR-PSQOSPRAVAVDKLAALLENHHNNH 526

RESULT 5
US-09-804-626-2
; Sequence 2, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
; APPLICANT: Lobel, Leslie
; APPLICANT: Lustbader, Joyce
; TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN OF
; FILE REFERENCE: 0575/62259/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/804,626
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-804-626-2

Query Match
Best Local Similarity 6.7%; Score 97; DB 10; Length 518;
Matches 60; Conservative 33; Mismatches 78; Indels 98; Gaps 13;

Oy 20 YGK-PEARMAAASGVSMFLYONGSEIADGRVVEVDIEVLKPNPSSFIITGKA 78
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 YNGCFEEYQSHAFNGTTLTSLKLEN-----VLEKEM--HNGAFR----- 364
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 79 GAOKSEKHNHNVSPAADQAFHTYGLMTPRYKVTVDGQEVKRTGEGOVSNLTGTGLRF 138
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 GA--TGPTLDTISYKRLALPSTGLE-----SIORLIAT----- 396
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; ORGANISM: Artificial Sequence

```

```

Oy 139 NLMSSESAAMVGOFDESKLPLFO-FINWVK-----VYKYPGQEGSGSDFTLDMT 187
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 -----SSYSLKILPSRFTFVNLEATLTLYSHCCAPRNLPTRKQ---NFSHSLIS 442
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 188 DNEFDITDGSRMKGGDMTFGCRNVYDLTDKNIYSRDKMLLALTRKGOESNGQVPRDDEPA 247
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 443 ENFSKOCSESTVVR-----VNNKTLYS--SMLAESELSGMDYEVGFCLPRTPCA 489
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 248 PNSSS-----VDKLAALLENHHNNH 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 PEPPAFNPCEIDMGVDKLAALLENHHNNH 518

RESULT 6
US-09-804-626-4
; Sequence 4, Application US/09804626
; Patent No. US20020128190A1
; GENERAL INFORMATION:
; APPLICANT: Lobel, Leslie
; APPLICANT: Lustbader, Joyce
; TITLE OF INVENTION: EXPRESSION OF PROPERLY FOLDED AND SOLUBLE EXTRACELLULAR DOMAIN
; FILE REFERENCE: 0575/62259/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/804,626
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-804-626-4

Query Match
Best Local Similarity 6.7%; Score 96.5; DB 10; Length 516;
Matches 26; Conservative 5; Mismatches 30; Indels 7; Gaps 2;

Oy 207 GNRVDTLDKNITSRDKMLLALTRKGOESNGQV-----PRDD--EPANSSSVKLA 259
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 449 GORSSLAEDNNESSYSGRFDMTYTERDYDLCEYVDYTGSPKPAENPCEDIMGVDKLA 508
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 260 LEHHNNH 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 509 LEHHNNH 516

RESULT 7
US-09-818-247-22
; Sequence 22, Application US/09818247
; Patent No. US20020102657A1
; GENERAL INFORMATION:
; APPLICANT: Mostov, Keith E.
; APPLICANT: Chapin, Steven J.
; APPLICANT: Richman-Eisenstat, Janice
; TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Secretory Componen
; FILE REFERENCE: 18062E-000910US
; CURRENT APPLICATION NUMBER: US/09/818,247
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: WO PCT/US01/09699
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,197
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,198
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence

```









```

RESULT 13
Q93UI1 PRELIMINARY; PRT; 205 AA.
ID Q93UI1
AC Q93UI1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Endo-1,3,1,4-beta-glucanase (Fragment).
DE Endo-1,3,1,4-beta-glucanase
DE uncultured bacterium.
OS Bacteria; environmental samples.
OC NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Mellado R.P.;
RT "Novel licheneses from soil.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF254960; AAK50613.1; -
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00732; Glyco_hydro_16; 1_16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT NON_TER
FT NON_TER 205
SQ SEQUENCE 205 AA; 22890 MW; E7CADAAFE02A3F51 CRC64;

Query Match 15.2%; Score 219; DB 2: Length 205;
Best Local Similarity 36.2%; Pred.No.5.1e-10;
Matches 55; Conservative 20; Mismatches 67; Indels 10; Gaps

QY 3 SAKDFSAGELTYLEEYQKFEARMKMAASGVSSNFLYQNGSEIADGRPWVEVDIEVL 62
DB 55 SYNKDCGENRSVQYCYGLYVRMPAKNTGIVSFFTYTGPT---GTPWDEIDIESL 111

QY 63 GKNPQSFQSNITKGAQKTSKHHAVSPADQAQPHHTYGLWTPNVRWTVDGQGVKRT 122
DB 112 GKDTTKVQFNYYTNGAG--NHEKLDADLGFDAANAYHMYAFADQPNKSIKWYDGO-LKHT 167

QY 123 EGGQVSNLTCQGLRFNLMSSESA-AWVGQFD 153
DB 168 ATTQVPAAPGK--IMNLNMGTCVDWLGSYN 197

RESULT 14
ID Q8U8N5 PRELIMINARY; PRT; 289 AA.
AC Q8U8N5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Endo-1,3,1,4-beta-glucanase
GN EXOK OR AT04055 OR AGR_L1600.
OC Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OC NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58";
RN RL Science 294:2317-2323(2001).
RN [2]

```

```
AC Q45691;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endo-beta-1,3-1,4-glucanase.
GN BEG1
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 8563;
RA van Rensburg P., van Zyl W.H., Pretorius I.S.;
RT "Over-expression of the Saccharomyces cerevisiae exo-beta-1,3-
RT glucanase gene together with the Butyrivibrio fibrisolvens endo-beta-
RT 1,4-glucanase gene and the Bacillus subtilis endo-beta-1,3-1,4-
RT glucanase gene in Saccharomyces cerevisiae.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60830; AAB05759.1; -.
DR HSSP; P27051; IGBG.
DR InterPro; IPR000757; Glyco_hydro.16.
DR Pfam; PF00722; Glyco_hydro.16; 1.
DR PRINTS; PF00737; GLHYDRASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
SQ SEQUENCE 239 AA; 26969 MW; C4219760D13F878A CRC64;

Query Match 16.3%; Score 234.5; DB 2; Length 239;
Best Local Similarity 33.7%; Pred. No. 3.5e-11;
Matches 58; Conservative 26; Mismatches 71; Indels 17; Gaps 7;

Qy 3 SAKDFSGAELVTLVEQYKGFTEARKMAAAGTSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db 80 SYNKPDCGENRSVQYGYGLYEVKMKPAKNTGIVSSFFTYTGPT- --GTPWDEIDIEFL 136

Qy 63 GKNPGSFSNIITGKAGAKTSEKHVAVSPA- --ADQAFHYTGLEWTPNYVWTVGQEV 119
Db 137 GKDTTKVQFNYYTNGA- --EENKLIADLGDAANAYHTYAFDQWPNSTKWYVDWQ-L 189

Qy 120 RKGTEGQVSLTGTGLRNFNLWSSESA- --WVGQFDESKLPLFQFINWVKYK 170
Db 190 KHTATQIPTAGK- --IMMNLWNGTGVDSWLGSYNGVN-PLYAHYDWVRYTK 238

RESULT 9
Q8RMP0
ID Q8RMP0 PRELIMINARY; PRT; 242 AA.
AC Q8RMP0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Beta-1,3-1,4-glucanase.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Sun J., Li W., Gu S., Xu Z., Zhao H., Xiao J.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490978; ANM08358.1; -.
SQ SEQUENCE 242 AA; 27238 MW; 285FED3FE76AE69A CRC64;

Query Match 16.2%; Score 232.5; DB 2; Length 242;
Best Local Similarity 34.9%; Pred. No. 5.2e-11;
Matches 59; Conservative 24; Mismatches 75; Indels 11; Gaps 6;

Qy 3 SAKDFSGAELVTLVEQYKGFTEARKMAAAGTSSMFLYQNGSEIADGRPWVEVDIEVL 62
Db 83 SYNKPDCGENRSVQYGYGLYEVKMKPAKNTGIVSSFFTYTGPT- --DGTWDEIDIEFL 139

Qy 63 GKNPGSFSNIITGKAGAKTSEKHVAVSPAADQAFHYTGLEWTPNYVWTVGQEVK 122
Db 137 GKDTTKVQFNYYTNGA- --EENKLIADLGDAANAYHTYAFDQWPNSTKWYVDWQ-L 189
```

```
Db 140 GKDTTKVQFNYYTNGAG- --NHXIVDLGDAANAYHTYAFDQWPNSTKWYVDGQ-LKHT 195
Qy 123 EGGQVSLTGTQGLRNFNLWSSESA- --AWVGQFDESKLPLFQFINWVKYK 170
Db 196 ATNQIPTPGK- --IMMNLWNGTGVDSWLGSYNGVN-PLYAHYDWVRYTK 241

RESULT 10
Q9FDC9
ID Q9FDC9 PRELIMINARY; PRT; 212 AA.
AC Q9FDC9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endo-1,3-1,4-beta-glucanase (Fragment).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Paenibacillaceae; Paenibacillus.
OX NCBI_TaxID=1406;
RN [1]
RP SEQUENCE FROM N.A.
RA Yao W., Wang Y., Song W., Yang K., Su Z.;
RT "Gene cloning of an antifungal protein.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF284449; AAG02415.1; -.
DR HSSP; P23904; IAJK.
DR InterPro; IPR000757; Glyco_hydro.16.
DR Pfam; PF00722; Glyco_hydro.16; 1.
DR PRINTS; PF00737; GLHYDRASE16.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
FT NON_TER 1
FT NON_TER 212
SQ SEQUENCE 212 AA; 24127 MW; 7E707E7PBEF7A440 CRC64;

Query Match 16.0%; Score 230; DB 2; Length 212;
Best Local Similarity 35.1%; Pred. No. 7e-11;
Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;

Qy 7 FSGAELVTLVEQYKGFTEARKMAAAGTSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66
Db 57 FDGGEYSKNTYRYGLYEVNKKPAKNTGIVSSFFTYTGPT- --ANGTQWDEIDIEFLGKDT 113

Qy 67 GSFSNIITGKAGAKTSEKHVAVSPAADQAFHYTGLEWTPNYVWTVGQEVRRTEGQ 126
Db 114 TKVQFNYYTNGTIGGH- --EKVVDLGLFDASSGFHTYAFDQWPGYIKWYVDG-VLKHT- --A 166

Qy 127 VSLTGTQGLRNFNLWSSESA- --AWVGQFDESKLPLFQFINWVKYK 172
Db 167 TTNIPKTPGQIMMNLWNGTGVDSWLGSYNGVN-PLYAEYDW- --KYT 210

RESULT 11
O07856
ID O07856 PRELIMINARY; PRT; 237 AA.
AC O07856;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta-(1,3-1,4)-glucanase precursor.
GN LICHENASE.
OS Streptococcus bovis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1315;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JBL;
RA Ekinci M., Flint H.J.;
RT "Isolation and overexpression of a gene encoding an extracellular
RT beta-(1,3-1,4)-glucanase from Streptococcus bovis JBL.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z92911; CAB07443.1; -.
DR HSSP; P23904; IAJK.
```

RN	[1]
RP	SEQUENCE FROM N.A.
RA	Rodriguez V., Mellado R.P.;
RT	"Novel lichenes from soil.";
RL	Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases .
DR	EMBL; AF254959; AAK50612.2 ; -
Dr	InterPro; IPR000757; Glyco_hydro_16.
Df	Pfam; PF00722; Glyco_hydro_16; 1.
Dg	PRINTS; PR00737; GLHYDRLASE16
DH	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_N_1.
FT	NON_TER
Ft	_
SQ	SEQUENCE      214 AA;    24175 MW;    3DBI92D2F6B4CAA3 CRC64;
Query Match                  16.8%; Score 241.5; DB 2; Length 214;	
Best Local Similarity         35.5% ; Pred. No. 8.4e-12;	
Matches	60; Conservative        24; Mismatches    74; Indels    11; Gaps
Qy	3 SAKDFSGAELYTL EEVYGKFEARKMAASAGTVSSMFLYONGSEIADGRPWVEVDIEVL 62     : ::    :    :  :    :  :    :  :    :
Db	55 SYNKFDCGENRSVTGYGLYEVRMKPAKTGIVSSFYTGTGPTE--GTPWDEIDIESL 111
Qy	63 GKNPGSFSNIITCKACAOKTSEKHHAVSPADAQAFTITYGLEWTPTNVVRTVDOGEVKR 121     :   :    :  :    :  :    :  :    :  :    :  :    :  :    :
Db	112 GKDTTKVOFNYYNTNGAG---NHKEKLADLGEDAAANAYHTAFDMOPNSIKWYVDQG-LKHT 167
Qy	123 EGGOVSNLTTQTGLRFNFWSSES-AWVGOFDESCLPLFOFINMWKVYK 170
Db	168 ATTQPPAAPGK--IMMNLNNGTGVDLDLSYNGVN-PLYAHYWVRVYRK 213
RESULT 7	
Q93GE6	PRELIMINARY; PRT; 214 AA.
ID	Q93GE6 PRELIMINARY; PRT; 214 AA.
AC	C93GG6;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DI	01-DEC-2002 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Endo 1-3,1-4-beta-glucanase (Fragment).
OS	uncultured bacterium.
OC	Bacteria; environmental samples.
OX	NCBL_Taxid=77133;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Rodriguez V., Mellado R.P.;
RT	"Novel lichenes from soil.";
RL	Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases .
DR	EMBL; AF254961; AAK50614.2 ; -
Dr	InterPro; IPR000757; Glyco_hydro_16.
Df	Pfam; PF00722; Glyco_hydro_16; 1.
Dg	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_N_1.
FT	NON TER
Ft	_
SQ	SEQUENCE      214 AA;    24131 MW;    LE9D8BF4300EBA5 CRC64;
Query Match                  16.4%; Score 235.5; DB 2; Length 214;	
Best Local Similarity         35.5% ; Pred. No. 2.5e-11;	
Matches	60; Conservative        23; Mismatches    75; Indels    11; Gaps
Qy	3 SAKDFSGAELYTL EEVYGKFEARKMAASAGTVSSMFLYONGSEIADGRPWVEVDIEVL 62     : ::    :    :  :    :  :    :  :    :
Db	55 SYNKFDCGENRSVTGYGLYEVRMKPAKTGIVSSFYTGTGPTE--GTPWDEIDIESL 111
Qy	63 GKNPGSFSNIITCKACAOKTSEKHHAVSPADAQAFTITYGLEWTPTNVVRTVDOGEVKR 122     :   :    :  :    :  :    :  :    :  :    :  :    :  :    :
Db	112 GKDTTKVOFNYYNTNGAG---NHKEKLADLGEDAAANAYHTAFDMOPNSIKWYVDQG-LKHT 167
Qy	123 EGGOVSNLTTQTGLRFNFWSSES-AWVGOFDESCLPLFOFINMWKVYK 170
Db	168 ATTQPPAAPGK--IMMNLNNGTGVDLDLSYNGVN-PLYAHYWVRVYRK 213
RESULT 8	
Q45691	PRELIMINARY; PRT; 219 AA.
ID	Q45691 PRELIMINARY; PRT; 219 AA.



GenCore version 5.1.3  
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 23.7951 Seconds  
 (without alignments)  
 2312.009 Million cell updates/sec

Title: US-09-654-652A-2  
 Perfect score: 1439  
 Sequence: 1 MVSADFGAELYTLEEVQY.....PNSSVDKLAALHHHHH 267

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL\_21.\*  
 1: sp\_archaea.\*  
 2: sp\_bacteria.\*  
 3: sp\_fungi.\*  
 4: sp\_human.\*  
 5: sp\_invertebrate.\*  
 6: sp\_mammal.\*  
 7: sp\_mhc.\*  
 8: sp\_organelle.\*  
 9: sp\_phage.\*  
 10: sp\_plant.\*  
 11: sp\_rodent.\*  
 12: sp\_virus.\*  
 13: sp\_vertebrate.\*  
 14: sp\_unclassified.\*  
 15: sp\_rvirus.\*  
 16: sp\_bacteriap.\*  
 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match %	ID	Description
1	266.5	18.5	851 16 Q9K7X5	Q9K7X5 bacillus ha
2	248.5	17.3	256 2 Q9APD8	Q9APD8 bacillus ci
3	247.5	17.2	276 2 Q45648	Q45648 bacillus sp
4	242.5	16.9	214 2 Q93GE8	Q93GE8 uncultured
5	242.5	16.9	802 2 Q9S310	Q9S310 ruminococcu
6	241.5	16.8	214 2 Q93GE7	Q93GE7 uncultured
7	235.5	16.4	214 2 Q93GE5	Q93GE5 uncultured
8	234.5	16.3	239 2 Q45691	Q45691 bacillus su
9	232.5	16.2	242 2 Q8RMP0	Q8RMP0 bacillus su
10	230	16.0	212 2 Q9FDC9	Q9FDC9 paenibacill
11	230	16.0	237 2 O07856	O07856 streptococc
12	221	15.4	246 16 Q97FD3	Q97FD3 clostridium
13	219	15.2	205 2 Q93U11	Q93U11 uncultured
14	206	14.3	289 16 Q8U8N5	Q8U8N5 agrobacteri
15	200	13.9	293 16 Q98C78	Q98C78 rhizobium l
16	190.5	13.2	302 16 Q9K7X6	Q9K7X6 bacillus ha

# SUMMARIES

## RESULT 1

Q9K7X5 ID Q9K7X5 PRELIMINARY; PRT; 851 AA.  
 AC Q9K7X5;  
 DT 01-OCT-2000 (TREMREL. 15, Created)  
 DT 01-OCT-2000 (TREMREL. 15, Last sequence update)  
 DE 01-DEC-2001 (TREMREL. 19, Last annotation update)  
 DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).  
 GN BGLS OR BH3232.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_taxid=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL; AP001518; BAB06951.1; -.  
 DR HSSP; P23904; 1AJK.  
 DR InterPro; IPR000757; Glyco\_hydro.16.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR Pfam; PF00722; Glyco\_hydro.16; 3.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR PRINTS; PR00737; GLYDRLASE16.  
 DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 3.  
 KW Hydrolase; Glycosidase; Complete proteome.  
 SQ SEQUENCE 851 AA; 98142 MW; C1C4F0F158400285 CRC64;

Query Match 18.5%; Score 266.5; DB 16; Length 851;  
 Best Local Similarity 29.7%; Pred. No. 5.1e-13;  
 Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;

QY 4 AKDFSGAELYTLEEVQYQKFEARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63





```
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145986.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.5%; Score 172.5; DB 21; Length 277;
Best Local Similarity 26.6%; Pred. No. 2.5e-06;
Matches 54; Conservative 34; Mismatches 74; Indels 41; Gaps 9;

Qy 24 NVSAKDFSCAELYTLEEVQYKFEARMKMAA--ASGTYSMFLYQNGSEIADGRPWVEVD 81
Db 48 SLSLDKSSGSGFQSNQEFLYGKAEVOMKLVPCNSAGTVTTFYLK-----SPGTTWDEID 101
Qy 82 IEVLGK---NPGSFQSNITGKAGAKTSEKHVAVSPAADOAFHTYGLWETPNYVRWTV 138
Db 102 FEFLGNISGHPYTLHTNVYT-KGSGDKEQQFHLWFDPTAN--FHTYCITWNQRIIFTVD 158
Qy 139 GOEVRKTEGGQVSNLTGTQGLRF-----NLWSSESAAWVG---QFDESKLPLFQF 185
Db 159 GPIREFMNAE-----SRGVFPPTKQPMRLYASLWEAEHMAWTRGGLKTDMSKAPETAY 212
Qy 186 IN-----WVKYKYTPGQGE 200
Db 213 YRNYNVEGCVWYNGKSVCPANSQ 235

Search completed: January 9, 2003, 12:16:20
Job time : 38.3542 secs
```



```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0150896.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

Query Match          9.5%; Score 173; DB 21; Length 282;
Best Local Similarity 25.6%; Pred No. 2.3e-06;
Matches 64; Conservative 36; Mismatches 92; Indels 58; Gaps 12;

QY 24 NVSAKDFSGAELYTLEEVOYGRFEARMKMAA--ASGTVSSMFLYQNGSIADGRPMVEYD 81
   ::|||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
Db 53 SLSLDKSSGSGFQSGNOEFLYKAEVQMKLVPCNSAGTVTFYLK-----SPGTTWDEID 106
   ::|||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
QY 82 IEVLCK---NPGSFOSNIITGKAGAKTSEKHHAVSPAADQAFHTYGLEWNTNYVRWTVD 138
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
Db 107 FEFLGNISGHPTLHTNVYT-KGTGDKEQOFLHFDPTVNV--FHTYCIITWNPQRIIFTVD 163
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
QY 139 GOEVRKTEGGQVSNL--TGTQGLRF--NLWSSESAAGV---QFDESKLPLFQFINWVKV 191
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
Db 164 GPIREFKNPEAIGVFPFTRQPMRLYASLWEAEHWAHTRGCGLEKTDWSKAPFTAF----- 217
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
QY 192 YKYTPGQEGGSDFTLDWTDNFDFTDGSRWGKGDWTFDGNRVLDLTKNIYSRDGLMLIAL 251
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
Db 218 YR-----NINV-DGCWANGKSSCSANSPPFTQK-----L 246
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
QY 252 TRKGOESFNG 261
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |
Db 247 DSNQGTMRMG 256
   |||:::| ||| |||::: |||:::| | | | | | | | | | | | | | | | | | | |

RESULT 14
ABB91456
ID ABB91456 standard; Protein; 282 AA.
XX
AC ABB91456;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 667.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
WPI; 2002-269010/31.
XX
DR
XX
PT Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -
XX
PS Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
(ABB90790-ABB94016) for herbicidally active compounds, comprising
aligning and comparing nucleic acid or amino acid sequences from plant
with nucleic acid or amino acid sequences from non-plant organisms using
suitable search parameters, where plant sequences having an E-value
greater by a factor of 3 than the E-value of most similar non-plant
```







CC protein accumulates in the ripe (but ungerminated) seeds. The  
 CC proportion of these to malted seeds is increased therefore, without  
 CC an unacceptable increase in viscosity.

SQ Sequence 237 AA;

Query Match 12.5%; Score 228; DB 11; Length 237;  
 Best Local Similarity 34.3%; Pred. NO. 4.4e-11;  
 Matches 60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;

QY 23 TNVSAKDESGAELTYLLEEVQYCKFEARMKMAAAGTSSMFLYQNGSEIADGRPWVEVDI 82  
 DB 75 TSSAYNKDFCAEYRTNITGYGLYEVSMKPAKNTGLVSSFFTYTGP---AHGTQWDEIDI 131  
 QY 83 EYLGKNGPSFQSNITITGKAGAKTSEKHAVSPAADQAFHTYGLEWTPNYVRWTVGQEV 142  
 DB 132 EPLGKDTTKVQFNYYTNGVGH---EKVISLGFDAKSGEHTYAFDPQGIYIKWYVDG-VL 187  
 QY 143 RKEGGQVSNLTGTQGLRFLNLSSESA-AWVGQFDESKLPLFQFINWVKVYKT 195  
 DB 188 KHT---ATANIPSTPGKIMMNLWNGTGVDWLGSYNGAN-PLYAEVDWV---KYT 235

RESULT 9  
 AAR06622  
 ID AAR06622 standard; protein; 237 AA.

AC AAR06622;

DT 09-JAN-1991 (first entry)

DE Hybrid (1,3-1,4)-pre-beta-glucanase.

KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.

OS Bacillus amyloliquefaciens, Bacillus macerans.

FH Key Location/Qualifiers  
 FT Domain 1..127  
 FT /label-amino-terminal half of B.macerans beta-glucanase  
 FT 131..234  
 FT /label-carboxyl-terminal half of B.amyloliquefaciens  
 FT beta-glucanase

PN W09009436-A.

PD 23-AUG-1990.

PF 16-FEB-1990; 90WO-DK00044.

PR 04-AUG-1989; 89DK-0003848.

PR 16-FEB-1989; 89DD-0325800.

XX (CARL-) CARLSBERG A/S.  
 XX (DEAK ) AKAD WISSENSCHAFT DDR.

XX Borriass R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;

DR WPI; 1990-275129/36.

DR N-PSDB; AAR05833.

XX New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid  
 PT gene obt'd. using Bacillus amyloliquefaciens and B.macerans genes

XX Disclosure; page 28; 84pp; English.

CC This hybrid protein is encoded by the beta-glucanase-H2 gene.  
 CC Following processing of the signal peptide the mature protein  
 CC is produced, comprising the amino terminus of the B.macerans  
 CC beta-glucanase and the carboxyl-terminal half of the B.amylo-  
 CC liquefaciens beta-glucanase. This hybrid protein is thermo-  
 CC stable and hydrolyses beta-glycosidic linkages in (1,3-1,4)-  
 CC beta-glucans. Reducing sugars are obt'd. at high temps. and

CC thus this enzyme can be used in the mfr. of food prods., esp.  
 CC beer and animal feed (e.g. for feeding poultry). See also  
 CC AAR05833.

SQ Sequence 237 AA;

Query Match 12.2%; Score 222.5; DB 11; Length 237;  
 Best Local Similarity 33.1%; Pred. NO. 1.3e-10;  
 Matches 57; Conservative 25; Mismatches 79; Indels 11; Gaps 6;

QY 23 TNVSAKDESGAELTYLLEEVQYCKFEARMKMAAAGTSSMFLYQNGSEIADGRPWVEVDI 82  
 DB 75 TSSAYNKDFCAEYRTNITGYGLYEVSMKPAKNTGLVSSFFTYTGP---AHGTQWDEIDI 131  
 QY 83 EYLGKNGPSFQSNITITGKAGAKTSEKHAVSPAADQAFHTYGLEWTPNYVRWTVGQEV 142  
 DB 132 EPLGKDTTKVQFNYYTNGAG---NHEKFADLGFDANAYHTYAFDQPNISIKWYDGO-L 187  
 QY 143 RKEGGQVSNLTGTQGLRFLNLSSESA-AWVGQFDESKLPLFQFINWVKVYK 193  
 DB 188 KHTATTQIPAAPGK--IMMNLWNGTGVDWLGSYNGVN-PIYAHYHMMRYK 236

RESULT 10  
 AAW37884  
 ID AAW37884 standard; Protein; 245 AA.

AC AAW37884;

DT 20-AUG-1998 (first entry)

DE Lichenase protein.

KW Lichenase; lica; fungus; enzyme; beta-1,4-glucan bond hydrolysis;  
 KW beta-1,3-linked glucan bond; grain-containing feed; grain treatment;  
 KW nutrient availability; brewing; fermentation.

OS Orpinomyces sp.

FH Key Location/Qualifiers  
 FT Peptide 1..29  
 FT /note= "signal peptide"  
 FT Protein 30..246  
 FT /note= "mature lichenase"

PN W09814595-A1.

XX 09-APR-1998.

PF 03-OCT-1997; 97WO-US17811.

PR 04-OCT-1996; 96US-0027882.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Chen H, Li X, Ljungdahl LG;

DR WPI; 1998-240094/21.

DR N-PSDB; AAV29067.

XX New isolated lichenase protein - is obtained from Orpinomyces PC-2,  
 PT used for treatment of grain to improve feeds or to improve brewing  
 PT and fermentation processes

XX Claim 1; Page 24-25; 41pp; English.

XX This sequence is the lichenase (lica) of Orpinomyces sp. strain PC-2  
 CC of the invention. The protein was purified from a fungus or a fungal  
 CC culture or from a recombinant DNA molecule having a fungal lichenase  
 CC coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan  
 CC bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave  
 CC beta-1,4-linked glucans. They can be used for the treatment of animal  
 CC grain-containing feeds to improve nutrient availability and for treatment



KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.  
XX  
OS Unidentified.  
XX  
PN WO200212511-A1.  
XX  
PD 14-FEB-2002.  
XX  
XX 27-JUL-2001; 2001WO-ES00303.  
XX  
XX 28-JUL-2000; 2000ES-0001922.  
PR  
XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Rodriguez Munoz V, Perez Mellado R;  
PI  
XX WPI; 2002-217195/27.  
DR N-PSDB; ABL53375.  
XX  
XX Producing nucleic acid encoding lichenase, useful for producing enzymes  
PT for improving filtration in brewing, comprises isolating the nucleic  
PT acid from soil bacteria  
XX  
XX Claim 18; Page 21-22; 27pp; Spanish.  
PS  
XX The present sequence is a protein sequence from a soil bacterium for an  
CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)  
CC activity. The enzyme is useful in brewing and for degrading  
CC beta-glucanases that cause problems during filtration.  
XX  
SQ Sequence 214 AA;  
Query Match 12.6%; Score 229; DB 23; Length 214;  
Best Local Similarity 36.1%; Pred No. 3.2e-11;  
Matches 56; Conservative 22; Mismatches 67; Indels 10; Gaps 5;  
QY 23 TNVSAKDFSGAELTYLEEVOYKGFARMKMAAAGTSSMFLYQNGSEIADGRPWVEVDI 82  
Db 60 TSPSYNKFDCGENRSVQTYGYGLYEVSMKPAKNTGIVSFFTYTGTE---GTPWDEIDI 116  
QY 83 EVLGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFHTYGLEWTPNYVWTVYDQGEV 142  
Db 117 EFLGKDTTKVQFNYYTNGAG---NHEKLDLGLFDAANAYHTYAFDMQPNKIKWYVDGQ-L 172  
QY 143 RKTEGGQVSNLTGTQGLRFLNLWSSESA-AWVGQFD 176  
Db 173 KHTATQIPAAFGK--IMNMLNNGTGVDDWLGSYN 205  
RESULT 7  
AAR03775  
ID AAR03775 standard; protein; 234 AA.  
XX  
AC AAR03775;  
XX  
DT 31-JUL-1990 (first entry)  
XX  
DE Thermostable beta-glucanase.  
XX  
XX Thermostable beta-glucanase; ss; Bacillus subtilis.  
KW  
XX Bacillus macerans.  
OS  
PN DD272102-A.  
XX  
PD 27-SEP-1989.  
XX  
PF 12-MAY-1988; 88DD-0315706.  
XX  
XX 12-MAY-1988; 88DD-0315706.  
PR  
XX (DEAK) AKAD WISSENSCHAFT DDR.  
PA  
XX

PI Borriess R;  
XX  
DR WPI; 1990-067913/10.  
DR N-PSDB; AAQ03519.  
XX  
PT Thermostable beta-glucanase production -  
PT using Bacillus subtilis transformed with gene from Bacillus  
PT macerans.  
XX  
PS Fig 1; ; 9pp; German..  
XX  
XX The gene encoding Bacillus macerans thermostable  
CC beta-glucanase is expressed in Bacillus subtilis. The  
CC enzyme is useful for lowering the viscosity of brewing  
CC mashers and in the production of feedstuff.  
XX  
SQ Sequence 234 AA;  
Query Match 12.5%; Score 228; DB 11; Length 234;  
Best Local Similarity 34.3%; Pred No. 4.3e-11;  
Matches 60; Conservative 25; Mismatches 74; Indels 16; Gaps 8;  
QY 23 TNVSAKDFSGAELTYLEEVOYKGFARMKMAAAGTSSMFLYQNGSEIADGRPWVEVDI 82  
Db 72 TSSAYNKFDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSFFTYTGTE---AHGTQWDEIDI 128  
QY 83 EVLGKNPGSFQSNITGKAGAKTSEKHHAVSPAADOAFHTYGLEWTPNYVWTVYDQGEV 142  
Db 129 EFLGKDTTKVQFNYYTNGVGH---EKVISLGLFDASKGFHTYAFDMQPGYIKWYVDG-VL 184  
QY 143 RKTEGGQVSNLTGTQGLRFLNLWSSESA-AWVGQFDESKLPLFQFINWVKYKYT 195  
Db 185 KHT---ATANIPSTPGKIMNMLNNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 232  
RESULT 8  
AAR05803  
ID AAR05803 standard; protein; 237 AA.  
XX  
AC AAR05803;  
XX  
DT 08-NOV-1990 (first entry)  
XX  
DE Heat-stable endo-beta-1,3-1,4-glucanase.  
XX  
KW Endo-beta-1,3-1,4-glucanase; barley; brewing.  
XX  
OS Bacillus macerans.  
XX  
PN DD275704-A.  
XX  
PD 31-JAN-1990.  
XX  
PF 23-SEP-1988; 88DD-0320082.  
XX  
PR 23-SEP-1988; 88DD-0320082.  
XX  
XX (DEAK) AKAD WISSENSCHAFT DDR.  
XX  
PI Borriess R, Wobus U, Mendel R-R, Baumlein H;  
XX  
DR WPI; 1990-210631/28.  
DR N-PSDB; AAQ05167.  
XX  
XX Prepn. of barley plants expressing heat stable beta-glucanase -  
PT by transforming cells with appropriate vector then regeneration  
PT giving seeds useful in brewing without conversion to malt.  
XX  
PS Disclosure; ; p; German.  
XX  
XX The corresp. gene and the plant promoter region are inserted into an  
CC expression vector and used to transform barley cells. The transformants  
CC are used to regenerate barley plants which are useful in brewing. This  
CC





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 36.3542 Seconds  
(without alignments)  
1279.206 Million cell updates/sec

Title: US-09-654-652a-3  
Perfect score: 1824  
Sequence: 1 MNIKKTAVKSALAAVAAAAA.....AKGAKVNPNGHKRYRVNFEH 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	257.5	14.1	308	AAW93001	B. alkalophilus be
2	250.5	13.7	242	AAAP95000	Bacillus subtilis
3	248	13.6	214	AAE07317	Barley recombinant
4	248	13.6	239	AAE06621	Hybrid (1,3-1,4)-p
5	229	12.6	208	ABB76858	Bacterial lichenas
6	229	12.6	214	ABB76859	Bacterial lichenas
7	228	12.5	234	AAE03775	Thermotable beta-
8	228	12.5	237	AAE03803	Heat-stable endo-b
9	222.5	12.2	237	AAE06622	Hybrid (1,3-1,4)-p
10	190.5	10.4	245	AAW37884	Lichenase protein.

11	185	10.1	282	21	AAG48467	Arabidopsis thalia
12	185	10.1	282	23	ABB93737	Herbicidally activ
13	173	9.5	282	21	AAK32464	Arabidopsis thalia
14	173	9.5	282	23	ABB91456	Herbicidally activ
15	172.5	9.5	277	21	AAK52114	Arabidopsis thalia
16	172.5	9.5	277	23	ABB93169	Herbicidally activ
17	170.5	9.3	269	21	AAG18644	Arabidopsis thalia
18	170.5	9.3	282	21	AAG18643	Arabidopsis thalia
19	170.5	9.3	282	23	ABB93168	Herbicidally activ
20	170.5	9.3	298	21	AAG18642	Arabidopsis thalia
21	164	9.0	286	23	ABB93103	Herbicidally activ
22	161	8.8	284	23	ABB93853	Herbicidally activ
23	159.5	8.7	287	23	ABB93104	Herbicidally activ
24	157.5	8.6	269	21	AAG20347	Arabidopsis thalia
25	157.5	8.6	269	21	AAG20350	Arabidopsis thalia
26	157.5	8.6	269	21	AAG52653	Arabidopsis thalia
27	157.5	8.6	269	23	ABB93167	Herbicidally activ
28	157.5	8.6	290	21	AAG20346	Arabidopsis thalia
29	157.5	8.6	297	21	AAG52652	Arabidopsis thalia
30	157.5	8.6	300	21	AAG20349	Arabidopsis thalia
31	156.5	8.6	269	21	AAG20329	Arabidopsis thalia
32	156.5	8.6	299	21	AAG20328	Arabidopsis thalia
33	155	8.5	845	22	AAAB9062	Clostridium perfr
34	154.5	8.5	284	23	ABB93854	Herbicidally activ
35	154.5	8.5	306	17	AAE97362	Oerskovia beta-1,3
36	154	8.4	247	18	AAW11593	Protein encoded by
37	154	8.4	247	20	AAE08308	A. thaliana merist
38	145.5	8.0	285	23	ABB93852	Herbicidally activ
39	145.5	8.0	307	21	AAG40361	Arabidopsis thalia
40	145.5	8.0	310	21	AAG40360	Arabidopsis thalia
41	145.5	8.0	310	23	ABB90942	Herbicidally activ
42	145.5	8.0	312	21	ABB40359	Arabidopsis thalia
43	145	7.9	263	18	AAW29455	Oerskovia xanthine
44	145	7.9	305	23	ABB91796	Herbicidally activ
45	144.5	7.9	307	21	AAG16385	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAW93001  
ID AAW93001 standard; Protein; 308 AA.  
XX  
AC AAW93001;  
XX  
DT 19-MAY-1999 (first entry)  
XX  
DE B. alkalophilus beta-glucanase protein.  
XX  
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;  
KW equipment; food industry; brewing.  
XX  
OS Bacillus alkalophilus.  
XX  
PN DE19732751-A1.  
XX  
PD 04-FEB-1999.  
XX  
PF 30-JUL-1997; 97DE-1032751.  
XX  
PR 30-JUL-1997; 97DE-1032751.  
XX  
PA (HENK ) HENKEL KGAA.  
XX  
PI Hillen W, Maurer K;  
XX  
DR WPI; 1999-122161/11.  
XX  
DR N-PSDB; AAX02912.  
XX  
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful  
PT for removing glucan and/or lichenan from membranes in the brewing  
PT industry





Query Match 5.2%; Score 95.5; DB 10; Length 1289;  
Best Local Similarity 23.6%; Pred. No. 7.3;  
Matches 59; Conservative 34; Mismatches 92; Indels 65; Gaps 13;

Db 1373 SSTNSDTSM 1381  
|: | |:

## RESULT 8

US-09-841-132-444

; Sequence 444, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 444

; LENGTH: 1770

; TYPE: PRT

; ORGANISM: Chlamydia trachomatis serovar D

US-09-841-132-444

## Query Match

Best Local Similarity 5.3%; Score 97; DB 10; Length 1770;

Matches 88; Conservative 46; Mismatches 172; Indels 104; Gaps 16;

QY 2 NIKKTAVKALAVAAAAAALTNNYSAKDFSGABLYTLEEYQYKGFARMKMAAASQTSS 61

Db 171 NIKSLLEVNIIVE-----KGGSVYAKERVSLN-----VTEATFSS 207

QY 62 MFLYQNGSEIADGRPWVEVDIEVLGNKPGSQNIITGRAGAKTSEKHHAVSPAADQAF 121

Db 208 -----NGGEQGGGGIYSEQDMLISDCNVHFOGNAAGATAVKOCLDEEMIVLLAECVDSL 262

QY 122 HTYGLWTPNYVMTVDGQE--VRKTEGGQVS----- 151

Db 263 SEDTLSTPETETESNGHQDQSEFEDTQVSESPSTSPDVLKGGGIYTEKSLTIT 322

QY 152 NLGTGGLRNLWSSBSAAWQGFDESKLPFLQFINWVXYKTPQGGEGGSDFTLTDWD 211

Db 323 GIGTIDFVSNI-ATDSGA--GVFTRENLSCTN-TNSLOFLKNSAQHGGAIVTQMSV 378

QY 212 NEDTFDGRWKG--GDWTFDGRNVDLTK-----NIYSRDGMILLALTRKGQESFNG 261

Db 379 TINTSEIITPPLIGVIFSEN----TAKHGCGGICTNKLISLSNLTVTLTNKAKEGG 434

QY 262 -----QVPRDDEPAQS--SSSAPASSSVPAS-----SSSVPASSSSSAFVPPSSS 305

Db 435 AIFTDIASIPITDTPSSSTPSSSPASTPEVVASAKINRFFASTAKPAAPS---LTEAES 491

QY 306 SATNAIGHMRTTPAVAKEHRLVNA-----KGAKVNPNGHKRYRYN 346

Db 492 DQDQDETSDTNSDIDVSIENILNVAINQNTSAKKGGAIVGKKAKLSRIN 541

## RESULT 9

US-09-801-368-362

; Sequence 362, Application US/09801368

; Patent No. US20020128250A1

; GENERAL INFORMATION:

; APPLICANT: Busby, Robert

; APPLICANT: Call, Brian

; APPLICANT: Hecht, Peter

; APPLICANT: Holtzman, Doug

; APPLICANT: Madden, Kevin

; APPLICANT: Maxon, Mary

; APPLICANT: Milne, Todd

; APPLICANT: No. US20020128250Alman, Thea

; APPLICANT: Royer, John

; APPLICANT: Salama, Sofie

; APPLICANT: Sherman, Amir  
; APPLICANT: Silva, Jeff  
; APPLICANT: Summers, Eric  
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
; FILE REFERENCE: 109272.147  
; CURRENT APPLICATION NUMBER: US/09/801,368  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 09/487,558  
; PRIOR FILING DATE: 2000-01-19  
; PRIOR APPLICATION NUMBER: US 60/160,587  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 362  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-362

Query Match 5.3%; Score 96.5; DB 10; Length 526;

Best Local Similarity 21.0%; Pred. No. 1.8;

Matches 79; Conservative 49; Mismatches 147; Indels 101; Gaps 15;

QY 36 YTLEEYQYKGFARMKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGNKPGSFQSN 95

Db 6 YTMQGYNLVLLKRLLEATA--RLEDVTIYQEG-----YIQNKLEASKNNKPS---- 51

QY 96 IITGKAGAKTSEKHHAVSPAADQ-----APHYTGLEWTPNVRW-----TVDGOEVRK 144

Db 52 --DSGADANTTNEPSAENAPEVEQDPKCTATFQSYICENIDPLVELSGKIDTVLVDALQL 109

QY 145 TEGGQVSNLTGTQGL--RFNLWSSESA-----WYQFDESKLPFLQFINWVK 190

Db 110 LKGGFQSQTLTFURAAVRSRKPDSQTADFASLRPINENIKLGQLKESN-----RQSK 162

QY 191 VVKYTPQGGEGGSDFTLTDN-----FDTFGSRWKGKDWI-----FDGNRVDLT 236

Db 163 YFAYLSALSEGAPLFWSWAVDTPVSMVTFDKDAAQF---WTNRIKLEYRSDPNAVEWV 218

QY 237 DKNIYSRDGMILL-----ALTRKGQESFNGQVPRDDEPAQSSSSAPASSSVPASS 288

Db 219 KKPLASFDNLKAYIKEYHTTGVSWKKDGMDFADAMAQSTKNTGATSSPASATAAPAP 278

QY 289 SSVPASSSSSAF-----VPPSSSSAT-----NATHGMRTTPAVAKEHNR-- 336

Db 279 PPPAPPASVFEISNDTPATSSDANKGGIGAVFAELNQGENITKGLKKVYDKSQOOTHNPE 338

QY 327 -----LVNAGAKVNP 337

Db 339 LRQSSVSVSTGSKSGP 354

## RESULT 10

US-09-841-132-195

; Sequence 195, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 195

; LENGTH: 821

; TYPE: PRT

; ORGANISM: Chlamydia

US-09-841-132-195



```
RESULT 6
US-09-815-242-13423
; Sequence 13423, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRN
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match          5.8%; Score 106.5; DB 10; Length 719;
Best Local Similarity 23.5%; Pred. No. 0.4;
Matches 76; Conservative 49; Mismatches 105; Indels 93; Gaps 20;

QY 57 GTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN-----PGSFQSNLITGKAGA 103
DB 417 GNITLQYALQQSRNV-----PAVET-LNKVGLNRAKTFNLGLIDIPSIHYSNAIS--SNT 469
QY 104 QKTSEKHHAVS---PAADQAFHTYGLEWTPNYVRWTV--DGOEVRKTEGGQVSNLTGTQG 158
DB 470 TESDKKYGASSEKMAAAYAFAFANGGTYKPMYIHKVVFSDGSE-----KEFSNV-GTRA 522
QY 159 LRENWSSSAWVQGFDESKLPF-----QFINVKVYKTPQOGE--GGSDFTLDWTD 211
DB 523 MK-----ETTAYM-MTDMKMTVLSYGTGRNAYLAWL-----POAGKTGTSNYTDEEIE 569
QY 212 NF-----DTPDG---SRGKGDWTFDGNRV-----DLTDKNIYSRDGMLLIA--- 250
DB 570 NHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLPLVNGLTVAAKVYRSMTYLSGGS 629
QY 251 -----LTKKGQESF-NGQVPRDDPAPQ-----SSSSAPASSSSSVPASSSSSV 295
DB 630 NPEDWNIPGELYNGEYFVKNGARSTWSSPAPQPPSTESSSSSDS3TSQSSSTTPSTN 689
QY 296 SFAFVPPSSSSATNAIHGMRTTP 318
DB 690 NSTTTNPNNTQOS-----NTTP 707

RESULT 7
US-09-738-363-2
; Sequence 2, Application US/09738363
```

```
; Patent No. US20010010932A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, Harry E.
; APPLICANT: Schwab, George E.
; APPLICANT: Payne, Jewel M.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Foncarrada, Luis
; TITLE OF INVENTION: Nematicidal Proteins
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/738,363
; FILING DATE: 15-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/076,137
; FILING DATE: 12-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-20CCCD3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; INDIVIDUAL ISOLATE: PS17
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC 1627) NRRL B-18651
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-738-363-2

Query Match          5.3%; Score 97.5; DB 10; Length 1385;
Best Local Similarity 23.7%; Pred. No. 5.5;
Matches 59; Conservative 36; Mismatches 91; Indels 63; Gaps 13;

QY 56 SCTVSSMFLYQNGSEIFA-----DGRPWVEVD-IEVLGKNPGSFQSNITGKAGAKTSEK 109
DB 1176 NGSIRSDISYQNIDAIVLPTLPKLRHWFMSDRFSEOGDTMAKFGQALNRAYAQLEOSTLL 1235
QY 110 HHAVSPAADQAFHTYGLEWTPNYVRWTVGQEVKRTGEGQVSNLTGTQGLRFLNWSSES 169
DB 1236 HNG-----HFTKDAANWTIEG-----DAHQITLEDGRRVRLRLPDWSS-- 1273
QY 170 ANVGFDESKLPFLFOFINWVKYKYT-PCQGGGGSDFTLD-----WTDNFTDFTG 218
DB 1274 --VSQMI-----IENFNDKEYNLVFGQGE--TVTLEHGETKYIETHHFAFNTT 1324
QY 219 SRWKGDWTFDGNRVDLTDKNIYSRDGMLI---LALTRKGQESFNGQVPRDDPAPQSSS 275
DB 1325 SO--ROGLTFESNKVTVT---ISSDEGEFLVDNIALV-----EAPLFTDDQNSEGNTA 1372
QY 276 SAPASSSV 284
```

```

Db 142 WLYSDFEDYSVANGETYSEIDVYELQDFWYEGHODDIYDMDLNLHAVYKENGQGVWKR 201
QY 103 ---AOKTSEKHHAVSPAADOAFHTYGLEWTPNYVRVTVDGOEVRKTEG---QOVSNLTGT 156
Db 202 KMYPQOLNKRWDPSKD--HHYCEVQNEIHHYVDGEVARKPNKYHWRPNMVTLS 259
QY 157 QGLR---FNLWSESAAWGQFDE-----SKLPLOFINWVKVYKTPGOGEGSDFTL 207
Db 260 LGLRKFVFFDNKNNAINPETDAKAREKLSOIPSTSMYDYVRVWEKSAGN----- 310
QY 208 DWTNDFTED-----GSRWKGDMFTDGNRVDLTKNIYSRDGLMLIALTRKQOESFN 260
Db 311 --TTNPPTSEVGLTKGSKLVLDHWDASTGITSVSN-----TKTGO--YA 354
QY 261 GOVPRDEPAQOSSAPASSSVPASSSVPASSSSAFVPPSS-----SSATN 309
Db 355 GSV-----NNASIAQIVTLKANTSYKVSFAFGKASSPGTSAYLGISKASN 398

```

```

RESULT 4
US-09-841-132-353
; Sequence 353, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Ajay
; APPLICANT: Bhatia, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 353
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-841-132-353

```

```

Query Match 5.9%; Score 108; DB 10; Length 583;
Best Local Similarity 22.4%; Pred. NO. 0.23;
Matches 89; Conservative 51; Mismatches 174; Indels 84; Gaps 19;

QY 8 VKSALAVAAAAALTNVSAKDFSGAELYTLEEVQYGFPEAKMKMAASGTVSSMFLYQN 67
Db 27 IQGAMIAQIKLPTVHIGPTAFLG-----LGVVDNNGNGARVQVWGSAPAAALGI-ST 80
QY 68 GSEI--ADGRP---WVEVDIEVLGNPGSFQSNIIITGKAGAKTSEKHHAVSPAADOAFH 122
Db 81 GDVITAVDGA PINSATAMADALNGHPGDVISVWTQTKSGGTRTGNVLAEGPPAE--FC 138
QY 123 TYGLEWTPNYVRVTVDGOEVRKTE-----GG-----QVSNLTGTQGLRFNL 163
Db 139 RYPSHWRP-----LDTQVSESPSTPSPDDVLGKGGGIYTEKSLTITGITIDFVSN 192
QY 164 WSESAAWVGQFDESKLPLOFINWVKVYKTPGOGEGSDFTLWNTDFTDGSRGWK 223
Db 193 -ATDSCA--GVFTKENLSCNT--TNSLQFLKNSAGQHGGAIVYQTMSTVNTTSEITPPP 248
QY 224 --GDWTFDGNRVDLTK-----NIYSRDGLMLIALTRKQOESFNG-----QVPRD 266
Db 249 LVGEVIFSEN---TAKHGCGGICINKLSLNLKVTVLTKNKSAKESGGAIFDGLASIPT 304
QY 267 DEPAQOS--SSAPASSSVPASSSVPASSSVPASSSVPASSSVPASSSVPASSSVPASS 317
Db 305 DTPESSTPSSSPASTPEVVASAKINRRFFASTAEPAAPS---LTEASDQTDQTTSDTN 361
QY 318 PAVAKEHRLVNA-----KGAIVPNHGKRVN 346
Db 362 SDIDVSIENILNVAINTQNTSARKKGAIIYKKAKLKSRIN 399

```

```

RESULT 5
US-09-765-272-2
; Sequence 2, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2

```

```

Query Match 5.9%; Score 107.5; DB 10; Length 666;
Best Local Similarity 23.5%; Pred. NO. 0.3;
Matches 76; Conservative 50; Mismatches 104; Indels 93; Gaps 20;

QY 57 GTVSSMELYONGSEADGRPWVEVDIEVLGKN-----PGSFQSNIIITGKAGA 103
Db 364 GNITLOYALQOSRNV----PAVET-LNKVGLNRKATFLNGLGIDYPSIHYSNALS--SNT 416
QY 104 OKTSEKHHAVS---PAADQAFHTYGLEWTPNYVRVTV--DGOEVRKTEGGGVSNLTGTQ 158
Db 417 TESDKKYCASSEKMAAAYAAAFANGCTYKPKYIHKVVPDQSE-----KEFSNV-GTRA 469
QY 159 LRFNLWSESAAWGQFDESKLPF-----QFINWVKVYKTPGQGE--GGSDFTLDWTD 211
Db 470 MK-----ETTAYM-MTDMKMTVLTGTGRNAYLAWL-----POAGKTGTSNYTDEIE 516
QY 212 NF-----DTFDG--SRWKGDMFTDGNRV-----DLTKNIYSRDGLMLILA--- 250
Db 517 NHIKTSQFVAPDELFAGYTRKYSMAVWTGYSNRLTPLVNGLTVAAYKYSRMVTYLSG 576
QY 251 -----LTRKQOESF--NGQVPRDDEPAQ-----SSSAPASSSVPASSSVPASSSV 295
Db 577 NPEDWNIEPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSSSDSSTSSSSTTPSTN 636
QY 296 SSFAVPPSSSSSATNAIHGMRTPT 318
Db 637 NSTTTNPNNTQOS-----NTTP 654

```

```

US-09-988-200-8
Query Match      6.3%; Score 114.5; DB 10; Length 545;
Best Local Similarity 20.6%; Pred. No. 0.06;
Matches 73; Conservative 61; Mismatches 131; Indels 89; Gaps 17;

Qy      8 VKSLAVAAAAAALTNNVSAKD---FSGAELTLEEVQYGRFEARKMA--RASGTVSSM 62
|| : : | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      82 VKISNGIAELTMRHNANNTPPDGGTYTSGIFKSYQKFTYGYFEAKIQAGIDGEGVCSF 141
|| : : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy      63 FLYQN-GSEIADGRP-VWEVDIEVL-----GKNPGSFQSNI-----ITGKAG----- 102

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:16 : Search time 8.88657 Seconds  
(without alignments)  
761.932 Million cell updates/sec

Title: US-09-654-652A-3  
Perfect score: 1824  
Sequence: 1 MNIKKTAVKSALAVAAAAA.....AKGAKVNPNGHKRYRVNFH 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 118974 seqs, 19401057 residues  
Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.pdb.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.pdb.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.pdb.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.pdb.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.pdb.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.pdb.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pdb.pdb.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.pdb.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.pdb.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pdb.pdb.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pdb.pdb.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pdb.pdb.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.pdb.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	6.6	287	10	US-09-734-569-152
2	117	6.4	620	10	US-09-988-200-6
3	114.5	6.3	545	10	US-09-988-200-8
4	108	5.9	583	10	US-09-841-132-353
5	107.5	5.9	666	10	US-09-765-272-2
-6	106.5	5.8	719	10	US-09-815-242-13423
7	97.5	5.3	1385	10	US-09-738-363-2
8	97	5.3	1770	10	US-09-841-132-444
9	96.5	5.3	526	10	US-09-801-368-362
10	96	5.3	821	10	US-09-841-132-195
11	96	5.3	1776	10	US-09-841-132-179
12	95.5	5.2	1289	10	US-09-738-363-4
13	93	5.1	156	9	US-09-925-301-1154
14	93	5.1	714	9	US-09-738-626-6377
15	91	5.0	279	9	US-09-738-626-3942
16	91	5.0	678	10	US-09-801-368-314
17	91	5.0	1751	10	US-09-841-132-445
18	91	5.0	1751	10	US-09-841-132-594
19	90	4.9	674	9	US-10-086-464-14

20	89.5	4.9	493	9	US-09-738-626-6780	Sequence 6780, Ap
21	89.5	4.9	1162	10	US-09-745-008-34	Sequence 34, Appl
22	89	4.9	499	12	US-10-052-586-592	Sequence 592, App
23	89	4.9	1367	10	US-09-801-368-108	Sequence 108, App
24	88.5	4.9	630	10	US-09-815-242-10036	Sequence 10036, A
25	88	4.8	1723	10	US-09-841-132-394	Sequence 394, App
26	88	4.8	1723	10	US-09-841-132-395	Sequence 395, App
27	87.5	4.8	467	10	US-09-863-547B-1	Sequence 1, Appli
28	87.5	4.8	2167	10	US-09-778-927A-61	Sequence 61, Appl
29	87.5	4.8	2201	12	US-10-100-912-2	Sequence 2, Appli
30	87	4.8	655	10	US-09-205-658-57	Sequence 57, Appl
31	87	4.8	655	10	US-09-844-353A-57	Sequence 57, Appl
32	87	4.8	1884	10	US-09-785-770A-17	Sequence 17, Appl
33	87	4.8	1907	10	US-09-785-770A-16	Sequence 16, Appl
34	86.5	4.7	765	10	US-09-975-326-4	Sequence 4, Appli
35	86.5	4.7	778	10	US-09-815-242-10728	Sequence 10728, A
36	86.5	4.7	1463	9	US-09-971-536-69	Sequence 69, Appl
37	86	4.7	498	10	US-09-925-302-475	Sequence 475, App
38	86	4.7	671	9	US-09-792-630-43	Sequence 43, Appl
39	86	4.7	671	9	US-10-080-376-43	Sequence 43, Appl
40	86	4.7	766	9	US-09-934-406-2	Sequence 2, Appli
41	86	4.7	766	10	US-09-975-326-2	Sequence 2, Appli
42	86	4.7	871	10	US-09-886-468-21	Sequence 21, Appl
43	85.5	4.7	341	12	US-10-072-152-4	Sequence 4, Appli
44	85.5	4.7	848	10	US-09-841-132-192	Sequence 192, App
45	85.5	4.7	1530	10	US-09-841-132-178	Sequence 178, App

ALIGNMENTS

RESULT 1  
US-09-734-569-152  
; Sequence 152, Application US/09734569  
; Patent No. US20020064816A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehrhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Cirpus, Petra  
; APPLICANT: Bischoff, Friedrich  
; APPLICANT: Frank, Markus  
; APPLICANT: Freund, Annette  
; APPLICANT: Duening, Elke  
; APPLICANT: Schmidt, Ralf-Michael  
; APPLICANT: Reski, Ralf  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv  
; FILE REFERENCE: BASF-NAE-1332-99-US  
; CURRENT APPLICATION NUMBER: US/09734,569  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 60/171,101  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 181  
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1  
; SEQ ID NO 152  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Physcomitrella patens  
US-09-734-569-152

Query Match	6.6%;	Score 120;	DB 10;	Length 287;
Best Local Similarity	23.0%;	Pred. No. 0.0089;		
Matches 49;	Conservative 30;	Mismatches 70;	Indels 64;	Gaps 11;
Oy	43	YGFTEARKMAA--ASGVTSWFLYQNGSEIADGRPWVEVDLEVLGKNPQS---FQSNII	97	
Db	69	YVDISAYIKMPFDSAGTIVTFYMSQDQ-----HYELDFLCNTSGQFLLHTNVF	122	
Oy	98	TKGAGAOKTSKHHAVSPAADAQAFHYTYGLEWTPNYRVYVDGQEVKTEGGQVSNLTGT-	156	
Db	123	VDGVGGRE-QQMYLCGDFPSAD--FHYFRFRSKDNVVFVDNKPVR-----VFKNLECTV	174	

```

; PRIOR APPLICATION NUMBER: 60/329,709
; PRIOR FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
PCT-US02-33165-3

```

```

Query Match      5.4%; Score 98; DB 1; Length 261;
Best Local Similarity 21.9%; Pred. No. 20;
Matches      70; Conservative 36; Mismatches 100; Indels 114; Gaps 17;

QY  52  MAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQSNLIITCKACAQKTSEKHH 111
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   15  LAQAPAMAEVOLLQSGAEV-----KKPGS-----SVQVSCK-- 46

QY  112 AVSPAADAQFHTYGLEWTPNYVRVTVDGQEVKRTGGQVSNLTGQGIJRFNLWSSESAAM 171
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   47  ----ASGTTSMYGFNW-----VRQAPG-----HGIE-----W 70

QY  172 VGFDESKLPFLPFINWVKYK-----YTPQGEG-----GSDFTLD 208
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   71  MGGI-----IPFGTSYAQKFRGRVTFADQATSTAYMELTNLRSDDTAVYYCARDGPD 126

QY  209 WTDNFDTFDGSRWCKGDWTFDGNKRVLDLTDKNIYSRDGMILIALTRKGQESFNGQ-----V 263
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   127 WEDG-DSYDGS--GRGFDFPWGQGLTVTVSSA-STKGPSVFPLAPSSKSTSGGTAALGCL 182

QY  264 PRDDEPAQSSS-SAPASSSV---PA--SSSSVPASSSSAFVPPSSSSATNATHGMRIT 317
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db   183 VKDYFPPEPTVSNWNSGALTSQVHTFFPAVLQSSGLYSLSSVTVFPSSSLGTQTYICNVNHK 242

QY  318 PAVAKEHRNLVNAKAKVNP 337

Db   243 PSNTKVDK-----KVEP 254

```

Search completed: January 9, 2003, 12:14:20  
Job time : 22.3495 secs





```

; LENGTH: 771
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-10-264-213-165

Query Match          5.7%; Score 103.5; DB 6; Length 771;
Best Local Similarity 22.7%; Pred. No. 28;
Matches 79; Conservative 37; Mismatches 117; Indels 115; Gaps 18;

QY 32 GAELTYLEEV-QYGFARMKMAAASGTVSSMFLYQNGSEIADGR----- 75
Db 493 GASVNTVEAGAYGAF-----ANGGYRYKPY-YVNRVVSADGNTOTFNSQGTAMKSS 544
QY 76 -PWVEVDI--EVLGNKPGS-----FOSNIITGKAGAKQKTSKHHAVSPAADQAFHTYG 125
Db 545 TAYMITDMLKGLVTKGTGTSAAISGLYA-----GKTGTFDYSDEELKNP-----ALNATG 596
QY 126 L---EWTNNYVR-----WTVDGQEVKRTGQVSNLTGTQGLRFLNLSSESAAWVGQFD 176
Db 597 IAKDAWFTGYTRNRVISVMT--GYDKPTSHGISVAEQTISQ----- 635
QY 177 ESKLPLFOFINVKYK----YTPGOGEGGSDFTL-DWTDNFDTFDGRWKGKGDWTFDGN 231
Db 636 -----KIYKALMSYT-SQNDNKKDWTKPDTVESYNILKGSNPGTATITGSAN 681
QY 232 RVDLTDKNIYSRQGLILALTRKGOESFNGQVPRDEPAPOSSSAPASSSSVPASSSSV 291
Db 682 ----TTKELYVR-----GHGPSHKAVAESSSSSSSSES-----SSSKE 717
QY 292 PASSSSAEPVPPSSSATWAIHGMRTTPAVAKEHRLNVLNAKAKVNPNG 339
Db 718 SSSSSNAVASIASHATSSSVSTAPSSAAPSKEETSSSSEPVSAG 765

RESULT 6
US-09-724-676-83341
; Sequence 83341, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
; LENGTH: 1554
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-83341

Query Match          5.7%; Score 103.5; DB 5; Length 1554;
Best Local Similarity 20.8%; Pred. No. 61;
Matches 46; Conservative 36; Mismatches 80; Indels 59; Gaps 8;

QY 35 LYTLEEVOY-QKFEARMKMAAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGNKPGSFQ 93
Db 1210 LTTIRAFRYEARFOOKLLEYTDSNNIASLFL-----TAANRWLEVRMEYIG----- 1255
QY 94 SNIITGKAGAOKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGQVSNL 153
Db 1256 ACVVLIAAVTSISNLSHRELSAGLVGLGLTYAL-MVSNYLNWVRNLADMEQLQGVAKRI 1314
QY 154 TGTQGLRFLNLSSESAAWVG-----OFDESKLPLPFOFINWV 189
Db 1315 HG-----LLKTEAESYEGLLAPSLIPKNWPDQKIQIQLNSVRYDSSLKPVKLKHNAL 1367
QY 190 KVKYKTPGQ-----GEGGSDFTLDWTFDGRWG 222
Db 1368 ----ISPQKIGICRTGSGKSSFLAFFRMTDFEGQLEG 1404

RESULT 7
US-09-724-676-83341
; Sequence 83341, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
; LENGTH: 1554
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-83341

Query Match          5.7%; Score 103.5; DB 5; Length 1554;
Best Local Similarity 20.8%; Pred. No. 61;
Matches 46; Conservative 36; Mismatches 80; Indels 59; Gaps 8;

QY 35 LYTLEEVOY-QKFEARMKMAAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGNKPGSFQ 93
Db 1210 LTTIRAFRYEARFOOKLLEYTDSNNIASLFL-----TAANRWLEVRMEYIG----- 1255
QY 94 SNIITGKAGAOKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGQVSNL 153
Db 1256 ACVVLIAAVTSISNLSHRELSAGLVGLGLTYAL-MVSNYLNWVRNLADMEQLQGVAKRI 1314
QY 154 TGTQGLRFLNLSSESAAWVG-----OFDESKLPLPFOFINWV 189
Db 1315 HG-----LLKTEAESYEGLLAPSLIPKNWPDQKIQIQLNSVRYDSSLKPVKLKHNAL 1367
QY 190 KVKYKTPGQ-----GEGGSDFTLDWTFDGRWG 222
Db 1368 ----ISPQKIGICRTGSGKSSFLAFFRMTDFEGQLEG 1404

RESULT 8
US-10-287-274-406
; Sequence 406, Application US/10287274
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THER
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287.274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-406

Query Match          5.4%; Score 99; DB 6; Length 571;
Best Local Similarity 21.8%; Pred. No. 40;
Matches 58; Conservative 34; Mismatches 88; Indels 86; Gaps 11;

QY 3 IKKTAVKSAK-----AVAAAA-----ALTTNWSAKDFSGAE-----LYT 37
Db 16 LKKSWSVTSISLILASGMAAFAAHAADDVKLKATKTNVAFSDFTPTTEYSTGKPNIIIVLT 75
QY 38 LEEVOYCKFEARMKMAAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGNKPGSFQSN-- 95
Db 76 MDDLYGQLP-----FDRGS-----FDPKTMENREV 101
QY 96 IITGKAGAOKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGQVSNLTG 155
```

```

US-09-724-676A-83341
; Sequence 83341, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83341
; LENGTH: 1554
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-83341

Query Match          5.7%; Score 103.5; DB 5; Length 1554;
Best Local Similarity 20.8%; Pred. No. 61;
Matches 46; Conservative 36; Mismatches 80; Indels 59; Gaps 8;

QY 35 LYTLEEVOY-QKFEARMKMAAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGNKPGSFQ 93
Db 1210 LTTIRAFRYEARFOOKLLEYTDSNNIASLFL-----TAANRWLEVRMEYIG----- 1255
QY 94 SNIITGKAGAOKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGQVSNL 153
Db 1256 ACVVLIAAVTSISNLSHRELSAGLVGLGLTYAL-MVSNYLNWVRNLADMEQLQGVAKRI 1314
QY 154 TGTQGLRFLNLSSESAAWVG-----OFDESKLPLPFOFINWV 189
Db 1315 HG-----LLKTEAESYEGLLAPSLIPKNWPDQKIQIQLNSVRYDSSLKPVKLKHNAL 1367
QY 190 KVKYKTPGQ-----GEGGSDFTLDWTFDGRWG 222
Db 1368 ----ISPQKIGICRTGSGKSSFLAFFRMTDFEGQLEG 1404

RESULT 8
US-10-287-274-406
; Sequence 406, Application US/10287274
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THER
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287.274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 406
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-406

Query Match          5.4%; Score 99; DB 6; Length 571;
Best Local Similarity 21.8%; Pred. No. 40;
Matches 58; Conservative 34; Mismatches 88; Indels 86; Gaps 11;

QY 3 IKKTAVKSAK-----AVAAAA-----ALTTNWSAKDFSGAE-----LYT 37
Db 16 LKKSWSVTSISLILASGMAAFAAHAADDVKLKATKTNVAFSDFTPTTEYSTGKPNIIIVLT 75
QY 38 LEEVOYCKFEARMKMAAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGNKPGSFQSN-- 95
Db 76 MDDLYGQLP-----FDRGS-----FDPKTMENREV 101
QY 96 IITGKAGAOKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGQVSNLTG 155
```







Search completed: January 9, 2003, 12:13:26  
Job time : 118.824 secs



Db 1282 AGVDIOLMEDFVLGVGSAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGOYSLG 1341  
QY 50 DGRPWVEVDIEVLGNPGSFQSNIIITGKAGACKTSEKHHAVSPAADQAPHTYGLEWTPNY 109  
Db 1342 ETQNDMKTRYGVLGESSASWTSRGVL-----ADALVEYRSLVGPVRET----FYALHFNPN-Y 1393  
QY 110 VR-----WTVDGQEVKRTGGQVSNLTGTOGLRFLNWSSESAAWVGQFDESKLPL 159  
Db 1394 VEVSYASMKFPFGFTGQGREARSFEDASLTNITPLCMKPEL-----AFIKGQFSE----- 1443  
QY 160 FOFIN-----WVKVYKTPGQG---EGGSDF-----TLDWTDNF 190  
Db 1444 ---VNSLGISYAW-EAYRKVEGGAVOLLLEAGFDWEGAPMDLPRQELRVALENNTWSSYF 1499  
QY 191 DTFDG 195  
Db 1500 STVLG 1504

Search completed: January 9, 2003, 12:14:40  
Job time : 7.79861 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 27.8125 Seconds  
(without alignments)  
1279.206 Million cell updates/sec

Title: US-09-654-652A-2  
Perfect score: 1439  
Sequence: 1 MVSARDFSGAEYLTLEVOY.....PNSSVDVKLAALAEHHHHH 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	257.5	17.9	308	AAW93001	B. alkalophilus be
2	246.5	17.1	242	AAW95000	Bacillus subtilis
3	244	17.0	214	AAE07317	Barley recombinant
4	244	17.0	239	AAE06621	Hybrid (1,3-1,4)-p
5	226	15.7	234	AAE03775	Thermotable beta-
6	226	15.7	237	AAE05803	Heat-stable endo-b
7	225	15.6	208	ABB76858	Bacterial lichenas
8	225	15.6	214	ABB76859	Bacterial lichenas
9	220.5	15.3	237	AAE06622	Hybrid (1,3-1,4)-p
10	189	13.1	245	AAW37884	Lichenase protein.

11	184	12.8	282	21	AAG48467	Arabidopsis thalia
12	184	12.8	282	23	ABB93737	Herbicidally activ
13	172	12.0	282	21	AAG32464	Arabidopsis thalia
14	172	12.0	282	23	ABB91456	Herbicidally activ
15	171.5	11.9	277	21	AAG52114	Arabidopsis thalia
16	171.5	11.9	277	23	ABB93169	Herbicidally activ
17	169.5	11.8	269	21	AAG18644	Arabidopsis thalia
18	169.5	11.8	282	21	AAG18643	Arabidopsis thalia
19	169.5	11.8	282	23	ABB93168	Herbicidally activ
20	169.5	11.8	298	21	AAG18642	Arabidopsis thalia
21	163	11.3	286	23	ABB93103	Herbicidally activ
22	161	11.2	284	23	ABB93853	Herbicidally activ
23	159	11.0	287	23	ABB93104	Herbicidally activ
24	154.5	10.7	284	23	ABB93854	Herbicidally activ
25	154	10.7	845	22	ABB98062	Clostridium perfr
26	150.5	10.5	269	21	AAG20347	Arabidopsis thalia
27	150.5	10.5	269	21	AAG20350	Arabidopsis thalia
28	150.5	10.5	269	21	AAG52653	Arabidopsis thalia
29	150.5	10.5	269	23	ABB93167	Herbicidally activ
30	150.5	10.5	290	21	AAG20346	Arabidopsis thalia
31	150.5	10.5	297	21	AAG52652	Arabidopsis thalia
32	150.5	10.5	300	21	AAG20349	Arabidopsis thalia
33	149.5	10.4	269	21	AAG20329	Arabidopsis thalia
34	149.5	10.4	299	21	AAG20328	Arabidopsis thalia
35	149.5	10.4	306	17	AAE97362	Oerskovia beta-1,3
36	147	10.2	247	18	AAW11593	Protein encoded by
37	147	10.2	247	20	AAV08308	A. thaliana merist
38	145.5	10.1	285	23	ABB93852	Herbicidally activ
39	145	10.1	263	18	AAW29455	Oerskovia xanthine
40	143	9.9	204	21	AAG48468	Arabidopsis thalia
41	143	9.9	435	18	AAW29456	Oerskovia xanthine
42	141	9.8	261	12	AAE11599	Beta-1,3-glucanase
43	140.5	9.8	303	18	AAW29457	Oerskovia xanthine
44	140.5	9.8	307	21	AAG40361	Arabidopsis thalia
45	140.5	9.8	310	21	AAG40360	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
AAW93001  
ID AAW93001 standard; Protein; 308 AA.  
XX  
AC AAW93001;  
XX  
DT 19-MAY-1999 (first entry)  
XX  
DE B. alkalophilus beta-glucanase protein.  
XX  
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;  
KW equipment; food industry; brewing.  
XX  
OS Bacillus alkalophilus.  
XX  
PN DE19732751-A1.  
XX  
PD 04-FEB-1999.  
XX  
PF 30-JUL-1997; 97DE-1032751.  
XX  
PR 30-JUL-1997; 97DE-1032751.  
XX  
PA (HENK) HENKEL KGAA.  
XX  
PI Hillen W, Maurer K;  
XX  
DR WPI; 1999-122161/11.  
DR N-PSDB; AA02912.  
XX  
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful  
PT for removing glucan and/or lichenan from membranes in the brewing  
PT industry

```

XX PS Claim 6; Page 5; 8pp; German.
XX CC This sequence represents a novel beta-glucanase isolated from Bacillus
CC alkalophilus DSM 9956. This enzyme is useful for removing glucan and/or
CC lichenan from membranes and equipment in the food industry, especially
CC the brewing industry.
XX SQ Sequence 308 AA;

Query Match 17.9%; Score 257.5; DB 20; Length 308;
Best Local Similarity 36.5%; Pred. No. 1.4e-16;
Matches 62; Conservative 22; Mismatches 73; Indels 13; Gaps 6;

QY 6 DFSGAEYLTLEEVQYQKFEARMKMAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65
DB 96 EYKAGELRTNQYQGLFEVNMKPAKSTGVSSLTFTY-TGPDWDNDPWEIDIEFLGKD 154
QY 66 PGFSFOSNIITKAGAQKTSEKHIAVSPADQAFHTYGLWTPNYVWTVGQEVKTEGG 125
DB 155 TTRVQFNFTNGVG---NNEHYHELGFDASEFNITYAFWRPESISWYVNGELVYTA--- 208
QY 126 QVSNLTGCT-QGLRFLNWSESA-AWVGQFDESKLPLFOFINNVKVKYKTP 173
DB 209 -TENIPQTPQKIMMNLPGIGVDGWTGTVFGDGEDTPVVTEDWV---RYTP 254

RESULT 2
AAP95000
ID AAP95000 standard; Protein; 242 AA.
XX
AC AAP95000;
DT 12-FEB-1997 (first entry)
XX
DE Bacillus subtilis lichenase.
XX
KW Beer production; fermentation; barley; beta-glucan; hydrolysis;
KW lichenase.
XX
OS Bacillus subtilis Y-25.
XX
PN JP01067181-A.
XX
PD 13-MAR-1989.
XX
PF 08-SEP-1987; 87JP-0224615.
XX
PR 08-SEP-1987; 87JP-0224615.
XX
PA (ASAK ) ASahi BREWERIES KK.
XX
DR WPI; 1989-119863/16.
DR N-P5DB; AAN95000.
XX
PT Recombinant plasmid used in beer prodn. - obt'd. by integrating
PT lichenase gene derived from Bacillus subtilis, into vector
PS Disclosure; Fig 4; 7pp; Japanese.
XX
CC The lichenase gene from Bacillus subtilis Y-25 is used for
CC transforming Bacillus hosts so that they show increased lichenase
CC expression. The recombinant lichenase enzyme produced by the
CC transformants is useful in beer production for decomposing beta-
CC glucan from barley.
XX SQ Sequence 242 AA;

Query Match 17.1%; Score 246.5; DB 10; Length 242;
Best Local Similarity 35.5%; Pred. No. 1.2e-15;
Matches 60; Conservative 24; Mismatches 74; Indels 11; Gaps 6;

QY 3 SAKDFSGAEYLTLEEVQYQKFEARMKMAASCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62

```

```

DB 83 SYNKFDCCGENRSVQYGYGLYEVNRMKPAKNTGIVSSFFTYTGPT---DGTPEIDIEFL 139
QY 63 GKNPGSFSOSNIITKAGAQKTSEKHIAVSPADQAFHTYGLWTPNYVWTVGQEVKRT 122
DB 140 GKDTTKVQFNFTNGAG---NHEKIVDLGFDAAAYHTYAFDQPNKSIKWYVDGO-LKHT 195
QY 123 EGGQVSNLTGTQGLRFLNWSESA-AWVGQFDESKLPLFOFINNVKVKYK 170
DB 196 ATNQIPTTPGK--IMNWLNGTGVDENLGSGVNV-PLYAHYDWMVRYTK 241

RESULT 3
AAE07317
ID AAE07317 standard; Protein; 214 AA.
XX
AC AAE07317;
XX
DT 06-NOV-2001 (first entry)
XX
DE Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.
XX
KW Barley; transgenic malt; carbohydrate degrading enzyme; chicken;
KW (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.
XX
OS Hordeum vulgare.
XX
PN W0200159141-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US04222.
XX
PR 10-FEB-2000; 2000US-0181473.
PR 09-NOV-2000; 2000US-0247126.
XX
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX
PI Von Wettstein D, Huang J, Horvath H;
XX
WPI; 2001-497082/54.
XX
PT New barley based foodstuff for animals, i.e. chicken comprising
PT recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -
PS Claim 8; Page 37-38; 43pp; English.
XX
CC The present invention relates to a food stuff comprising barley feed and
CC transgenic barley malt where the transgenic barley malt comprises a
CC recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-
CC glucanase. The food stuff is useful to feed animal, preferably chickens
CC Barley is cheaper than corn, the principal foodstuff of chickens.
CC Chickens cannot efficiently utilise barley as an energy source as they
CC do not possess a gut enzyme that depolymerises beta-D-glucan, major
CC carbohydrate present in the barley endosperm. The invention provides a
CC barley based foodstuff which comprises a recombinant carbohydrate
CC degrading enzyme improving the nutritional value of the foodstuff.
CC The present sequence is barley recombinant, thermostable
CC (1,3-1,4)-beta-glucanase enzyme.
XX SQ Sequence 214 AA;

Query Match 17.0%; Score 244; DB 22; Length 214;
Best Local Similarity 35.5%; Pred. No. 1.7e-15;
Matches 61; Conservative 26; Mismatches 69; Indels 16; Gaps 8;

QY 3 SAKDFSGAEYLTLEEVQYQKFEARMKMAASCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62
DB 55 SYNKFDCCGENRSVQYGYGLYEVNRMKPAKNTGIVSSFFTYTGPT---GTPWEIDIEFL 111
QY 63 GKNPGSFSOSNIITKAGAQKTSEKHIAVSPADQAFHTYGLWTPNYVWTVGQEVKRT 122
DB 112 GKDTTKVQFNFTNGVGGH---EKVISLGFDAKGFHTYAFDQPGYIKWYVDG-VLKHT 167

```



Db 168 ---ATANIPSTPGKIMNWLNGTGVDDWLGSYNGAN-PLYAEDWV---K\* 212

AAR06621  
ID AAR06621 standard; protein; 239 AA.

XX  
DT 09-JAN-19XX  
DE Hybrid (1,3-1,4)-pre-beta-

XX Hybrid pre-beta-glucanase; glucans;  
KW

OS Bacillus amyloliquefaciens, Bacillus macerans.

Key	Location/Qualifiers
FH	1
FM	120

FT	/label=amino termin
FT	133 236
	Domain

FT  
XX  
/label=carboxyl-terminal of B.macerans

PN W09009436-A.  
XX

FD 23-AUG-1990.  
XX

XX  
XX

PR 16-FEB-1989; 89DD-0325800.  
YY

PA (CARL-) CARLSBERG A/S.  
PA (DEAK ) AKAD WISSENSCHAET DD

XX PI Borriß R, Hofemeister J, Th

XX  
DR WPI; 1990-275129/36.

DK N FSDS, AAQ05052.  
XX

gene obtd. using Bacillus am

PS Disclosure; page 26; 84pp; English.  
yy

CC This hybrid protein is encoded by t  
CC Following processing of the signal

CC is produced, comprising the amino terminus of the amyloid- $\beta$  (A $\beta$ ) and the carboxyl-terminal half of the A $\beta$ . A $\beta$  is produced, comprising the amino terminus of the amyloid- $\beta$  (A $\beta$ ) and the carboxyl-terminal half of the A $\beta$ . A $\beta$  is produced, comprising the amino terminus of the amyloid- $\beta$  (A $\beta$ ) and the carboxyl-terminal half of the A $\beta$ .

CC and hydrolyses beta-glycosidic linkages in (1,3-1,4)-beta-glucosylated B.macerans beta-glucanase. This hybrid protein is thermostable

reducing sugars are added, at high temps. and thus this enzyme can be used in the mfr. of food prods., esp. beer and animal feed (eg

XX XX  
US  
SOLUTIONS 230 000

Query Match

Best Local Similarity 35.5%, Pred. No. 2e-15;  
Matches 61; Conservative 26; Mismatches 69; Indels 1

QY 3 SAKDFSGAELYTLEEVQYCKFEAPMKMAAASGTVSSMFLYONGSEIADGRPWVEVDIEVL 62

Db 80 SYNKFDGGRNSVQTYGYGLYEVNRMKPAKNTGIVSSFFTTGPTGTE--GTPWDEIDIEFL 136

QY 63 GKNPGSFQSNII TGKAGAOKTSEKHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKKT 122

Db 137 GKDTTKVQFNYYTNGVGGH---EKVISLGFDAKSGFHTYAFDWQPGYIKWYVDG-VLKHT 192

QY- 123 EGGQVSNLTGTQG-LRFNLWSSES-AWVGQFDESKLPFQFINWVKVYKYT 172

KW Endo-beta-1,3-1,4-glucanase; barley; brewing.  
 XX  
 OS Bacillus macerans.  
 XX  
 PN DD275704-A.  
 XX  
 PD 31-JAN-1990.  
 XX  
 XX 23-SEP-1988; 88DD-0320082.  
 PF  
 XX 23-SEP-1988; 88DD-0320082.  
 PR  
 XX (DEAK ) AKAD WISSENSCHAFT DDR.  
 PA  
 XX Borriass R, Wobus U, Mendel R-R, Baumlein H;  
 PI  
 XX WPI: 1990-210631/28.  
 DR  
 DR N-PSDB; AAO05167.  
 XX  
 XX Prepn. of barley plants expressing heat stable beta-glucanase -  
 PT by transforming cells with appropriate vector then regeneration  
 PT giving seeds useful in brewing without conversion to malt.  
 PT  
 XX Disclosure: ; p; German.  
 XX  
 CC The corresp. gene and the plant promoter region are inserted into an  
 CC expression vector and used to transform barley cells. The transformants  
 CC are used to regenerate barley plants which are useful in brewing. This  
 CC protein accumulates in the ripe (but ungerminated) seeds. The  
 CC proportion of these to malted seeds is increased therefore, without  
 CC an unacceptable increase in viscosity.  
 XX  
 SQ Sequence 237 AA;  
 Query Match 15.7%; Score 226; DB 11; Length 237;  
 Best Local Similarity 35.1%; Pred. No. 1.1e-13;  
 Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;  
 QY 7 FSGAELTYLEEVQYKGFARMKMAAAGCTVSSMFLYQNGSEIADGRPWVVDIEVLGKNP 66  
 DB 82 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSFFTYTGP---AHGTQWDEIDIEFLGKDT 138  
 QY 67 GSFQSNITGKAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKTEGGQ 126  
 DB 139 TRVQFNYYTNGVGGH---EKVISLGFDAKSGPHTYAFDMQPGYIKWYVDG-VLKHT---A 191  
 QY 127 VSNLGTQGLRFNLWSSESA-AWVGQFDESKLPLFQFINWVKYKYT 172  
 DB 192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWV---KYT 235  
 RESULT 7  
 ABB76858  
 ID ABB76858 standard; Protein; 208 AA.  
 XX  
 AC ABB76858;  
 XX  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Bacterial lichenase #1.  
 XX  
 KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200212511-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 27-JUL-2001; 2001WO-ES00303.  
 XX  
 PR 28-JUL-2000; 2000ES-0001922.  
 XX  
 PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PI Rodriguez Munoz V, Perez Mellado R;  
 XX  
 XX WPI: 2002-217195/27.  
 DR N-PSDB; ABL53375.  
 XX  
 PT Producing nucleic acid encoding lichenase, useful for producing enzymes  
 PT for improving filtration in brewing, comprises isolating the nucleic  
 PT acid from soil bacteria -  
 XX  
 XX Claim 18; Page 21-22; 27pp; Spanish.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an

PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX  
 PI Rodriguez Munoz V, Perez Mellado R;  
 XX  
 DR WPI: 2002-217195/27.  
 DR N-PSDB; ABL53374.  
 XX  
 PT Producing nucleic acid encoding lichenase, useful for producing enzymes  
 PT for improving filtration in brewing, comprises isolating the nucleic  
 PT acid from soil bacteria -  
 XX  
 PS Claim 18; Page 18-19; 27pp; Spanish.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an  
 CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)  
 CC activity. The enzyme is useful in brewing and for degrading  
 CC beta-glucanases that cause problems during filtration.  
 XX  
 SQ Sequence 208 AA;  
 Query Match 15.6%; Score 225; DB 23; Length 208;  
 Best Local Similarity 36.2%; Pred. No. 1.2e-13;  
 Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;  
 QY 3 SAKDFSGAELTYLEEVQYKGFARMKMAAAGCTVSSMFLYQNGSEIADGRPWVVDIEVL 62  
 DB 58 SYNKFDCCGENRSVQTYGYGLYEVKMKPAKNTGIVSFFTYTGPTE---GTPWDEIDIEFL 114  
 QY 63 GKNPQSFQSNITGKAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKRT 122  
 DB 115 GKDTTKVQFNYYTNGAG---NHEKLDLGLFDAANAYHTYAFDMQPNISIKWYVDGQ-LKHT 170  
 QY 123 EGGQVSNLTGTQGLRFNLWSSESA-AWVGQFD 153  
 DB 171 ATTQIPAPAGK--IMMNLWNGTGVDDWLGSYN 200  
 RESULT 8  
 ABB76859  
 ID ABB76859 standard; Protein; 214 AA.  
 XX  
 AC ABB76859;  
 XX  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Bacterial lichenase #2.  
 XX  
 KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200212511-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 27-JUL-2001; 2001WO-ES00303.  
 XX  
 PR 28-JUL-2000; 2000ES-0001922.  
 XX  
 PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PI Rodriguez Munoz V, Perez Mellado R;  
 XX  
 XX WPI: 2002-217195/27.  
 DR N-PSDB; ABL53375.  
 XX  
 PT Producing nucleic acid encoding lichenase, useful for producing enzymes  
 PT for improving filtration in brewing, comprises isolating the nucleic  
 PT acid from soil bacteria -  
 XX  
 XX Claim 18; Page 21-22; 27pp; Spanish.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an

CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)  
 CC activity. The enzyme is useful in brewing and for degrading  
 CC beta-glucanases that cause problems during filtration.  
 XX  
 SQ Sequence 214 AA;

Query Match 15.6%; Score 225; DB 23; Length 214;  
 Best Local Similarity 36.2%; Pred. No. 1.2e-13;  
 Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;

QY 3 SAKDFSGAELYLTLEEVQYGFKFEARMKMAASCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62  
 DB 63 SYNKDCGGRNSVQYGYGLYEVMRKPAKNTGIVSFFTYTGP---GTPWDEIDIEFL 119  
 QY 63 GRNPGSFQNIITGKAGAKTSEKHHAVSPAADQAPHTYGLEWTPNYVWTVDGQEVKRT 122  
 DB 120 GRDTRVQFNYYTNGAG---NHEKLDLGFDAANAYHTYAFDMQPSIKWYVDGQ-LKHT 175  
 QY 123 EGQVSNLTGTGGLRPNLWSSSA-AWVGQFD 153  
 DB 175 ATTQIPAPGK--IMNMLWNGTGVDWLGSYN 205

RESULT 9  
 AAR06622  
 ID AAR06622 standard; protein; 237 AA.  
 XX  
 AC AAR06622;  
 DT 09-JAN-1991 (first entry)  
 DE Hybrid (1.3-1.4)-pre-beta-glucanase.  
 XX  
 KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.  
 XX  
 OS Bacillus amyloliquefaciens, Bacillus macerans.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..127  
 FT /label=amino-terminal half of B.macerans beta-glucanase  
 FT 131..234  
 FT /label=carboxyl-terminal half of B.amyloliquefaciens  
 FT beta-glucanase  
 XX  
 PN W09009436-A.  
 XX  
 PD 23-AUG-1990.  
 XX  
 PF 16-FEB-1990; 90WO-DK00044.  
 XX  
 PR 04-AUG-1989; 89DK-0003848.  
 PR 16-FEB-1989; 89DD-0325800.  
 XX  
 PA (CARL-) CARLSBERG A/S.  
 PA (DEAK ) AKAD WISSENSCHAFT DDR.  
 XX  
 PI Borriass R, Hofemeister J, Thomsen KK, Olsen O, Vonwettstein D;  
 XX  
 XX WPI: 1990-275129/36.  
 DR N-PSDB; AAR05833.  
 XX  
 PT New thermostable (1.3-1.4)-beta-glucanase - prepd. using hybrid  
 PT gene obt'd. using Bacillus amyloliquefaciens and B.macerans genes  
 XX  
 PS Disclosure; page 28; 84pp; English.  
 XX  
 CC This hybrid protein is encoded by the beta-glucanase-H2 gene.  
 CC Following processing of the signal peptide the mature protein  
 CC is produced, comprising the amino terminus of the B.macerans  
 CC beta-glucanase and the carboxyl-terminal half of the B.amylo-  
 CC liquefaciens beta-glucanase. This hybrid protein is thermo-  
 CC stable and hydrolyses beta-glycosidic linkages in (1.3-1.4)-  
 CC beta-glucans. Reducing sugars are obt'd. at high temps. and

CC thus this enzyme can be used in the mfr. of food prods., esp.  
 CC beer and animal feed (e.g. for feeding poultry). See also  
 CC AAR05833.  
 XX  
 SQ Sequence 237 AA;

Query Match 15.3%; Score 220.5; DB 11; Length 237;  
 Best Local Similarity 33.9%; Pred. No. 3.9e-13;  
 Matches 56; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

QY 7 FSGAELYLTLEEVQYGFKFEARMKMAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66  
 DB 82 FPCAERYSTNIYGYGLYEVSMKPAKNTGIVSFFTYTGP---AHGTQWDEIDIEFLGKDT 138  
 QY 67 GSFQSNLIITGKAGAKTSEKHHAVSPAADQAPHTYGLEWTPNYVWTVDGQEVKRTGEGQ 126  
 DB 139 TKVQFNYYTNGAG---NHEKFADLGFDAANAYHTYAFDMQPSIKWYVDGQ-LKHTATTQ 194  
 QY 127 VSNLTGTGGLRPNLWSSSA-AWVGQFDESKLPFLQFINWVYK 170  
 DB 195 IPAAPGK--IMNMLWNGTGVDWLGSYNGVN-PIAHYHMMRYK 236

RESULT 10  
 AAW37884  
 ID AAW37884 standard; protein; 245 AA.  
 XX  
 AC AAW37884;  
 DT 20-AUG-1998 (first entry)  
 DE Lichenase protein.  
 XX  
 KW Lichenase; licA; fungus; enzyme; beta-1,4-glucan bond hydrolysis;  
 KW beta-1,3-linked glucan bond; grain-containing feed; grain treatment;  
 KW nutrient availability; brewing; fermentation.  
 XX  
 OS Orpinomyces sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..29  
 FT /note= "signal peptide"  
 FT Protein 30..246  
 FT /note= "mature lichenase"  
 XX  
 PN W09814595-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 03-OCT-1997; 97WO-US17811.  
 XX  
 PR 04-OCT-1996; 96US-0027882.  
 XX  
 PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
 XX  
 PI Chen H, Li X, Ljungdahl LG;  
 XX  
 DR WPI: 1998-240094/21.  
 DR N-PSDB; AAV29067.  
 XX  
 PT New isolated lichenase protein - is obtained from Orpinomyces PC-2,  
 PT used for treatment of grain to improve feeds or to improve brewing  
 PT and fermentation processes  
 XX  
 PS Claim 1; Page 24-25; 41pp; English.  
 XX  
 CC This sequence is the lichenase (licA) of Orpinomyces sp. strain PC-2  
 CC of the invention. The protein was purified from a fungus or a fungal  
 CC culture or from a recombinant DNA molecule having a fungal lichenase  
 CC coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan  
 CC bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave  
 CC beta-1,4-linked glucans. They can be used for the treatment of animal  
 CC grain-containing feeds to improve nutrient availability and for treatment

```
CC of grain (e.g. barley or wheat) in the brewing and fermentation
CC industries to increase carbon substrate availability and to maximise
CC production of desired products.
XX
SQ Sequence 245 AA;

Query Match 13.1%; Score 189; DB 19; Length 245;
Best Local Similarity 32.9%; Pred. No. 4.7e-10;
Matches 56; Conservative 16; Mismatches 76; Indels 22; Gaps 8;

QY 5 KDFSGAELTLEVQ-----YKPEARKMAAASGTVSSMFLYONGSEIADGRPWVEVD 58
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 RDGSG---YTCGEYRTKNYGYGMFQVNMKPIKPGVSSFFTYTGPS---DGTKWDETD 136
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 IEVLGNPCGSFOSNIITKAGAKTSEKHVAVSPAADQAFHTYGLEWTPNYVYRWTVDGQE 118
   || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 IEFLGYDTTKVFNYYTNCQGH--EHLHYLGFASQGFHYGFWRNSITWYVDGTA 193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 119 VRKTEGGQVSNLTGTG-LRFNLWSESA-AWVGQFDESKLPLFOFINWV 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 194 VVTA----YDNIPDTPGKIMMNAWNGIGYDDMLRPFN-GRTNISAYYDWNV 238
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
AAG48467
ID AAG48467 standard; Protein; 282 AA.
XX
AC AAG48467;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61207.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134769.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
```

```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.

Query Match 12.8%; Score 184; DB 21; Length 282;
Best Local Similarity 27.6%; Pred. No. 1.8e-09;
Matches 59; Conservative 33; Mismatches 98; Indels 24; Gaps 9;

QY 2 VSAKDFSGAELYLEVQYCKFEARKMAA--ASGTVSSMFLYONGSEIADGRPWVEVDI 59
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
Db 54 LSLDKFSGSGFQSHOFLYCKVEQMKLVPGNSAGTVTFYLK-----SPGTTWDEIDF 107
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
QY 60 EVLGK---NPGSFQSNITTKAGAKOITSEKHAVSPAADQAPHTYGLEWTPNVYRVTVDG 116
|||: :| |||: :| |||: :|||: :| |||: :| |||:
Db 108 EFLGNISGHPYTLHTNVYT-KGTGDKQQFHLWFDPTVD--FHTYCIWNQPVITIDG 164
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
QY 117 QEVKRTGGQVSNL--TGTQGLRF--NLWSSESAAMWG---QFDESKLPFLQFINNVKY 169
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
Db 165 IPIREFKNEALGVPPFKHPMRLYASLWEAEHWATRGGLEKTDWSKAPTAFYRNVD 224
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
QY 170 KYTPGQGGSGDFTLDWTDNFDFTDGS---RWGK 200
:| |||: :| ||| | |||: :|||: :| |||: :| |||:
Db 225 ACVWSNKGKSCSANSWFTQVLDFKGNRYKWAQ 258
:| |||: :| ||| | |||: :|||: :| |||: :| |||:

RESULT 12
ABB93737
ID ABB93737 standard; Protein; 282 AA.
XX
AC ABB93737;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2948.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -
XX
PS Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
(ABB90790-ABB94016) for herbicidally active compounds, comprising
aligning and comparing nucleic acid or amino acid sequences from plant
with nucleic acid or amino acid sequences from non-plant organisms using
suitable search parameters, where plant sequences having an E-value
greater by a factor of 3 than the E-value of most similar non-plant
sequences are selected. The polypeptides or nucleic acids encoding them
are useful for identifying modulators. The identified modulators are
useful as herbicides.
XX
```

[illegible]

```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.0%; Score 172; DB 21; Length 282;
Best Local Similarity 25.9%; Pred. No. 2.6e-08;
Matches 63; Conservative 34; Mismatches 88; Indels 58; Gaps 12;

QY 8 SGAELYTLEEVOYKFEARKMAA--ASCTVSSMFLYQNGSEIADGRPWVVDIEVLGK- 64
||: :| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 60 SGSGFQSNQOEFLYKAEVQMKLVPGNSAGTVTFYLK-----SPGTTWDEIDFEFLGNI 113
QY 65 --NPGSFOSNIITGKAGAKTSEKHAYSPAADQAFHTYGLEWTPNYVVTVDGQEVKRT 122
||: :| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 114 SGHPYTLHTNVTY-KGTGDKQOQFHLWFDPTVN--FHTYCIITWNQRIIFTVDGPIREF 170
QY 123 EGGQVSNL--TGTOGLRF--NLWSSESAAWG---QFDESKLPFLQFINWVKVYKYPGQ 175
||: :| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 171 KNPEAIGVFPFTRQPMRLYASLWEAEHWATRGLEKTDWSKAPPTAF-----YR----- 219
QY 176 GEGGSDFTLDWTDNFDFTDGSRWGKGDTFDCGNRVLDLTDKNIYSRDGMILALTRKGQES 235
||: :| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 220 -----NYNV-DGCVWANGKSSCSANSPWFTQK-----LDSNGQTR 253

QY 236 PNG 238
Db 254 MKG 256

RESULT 14
ABB91456
ID ABB91456 standard; Protein; 282 AA.
XX
AC ABB91456;
XX
DT 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 667.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -
PS Claim 5; SEQ ID NO 667; 261pp + Sequence Listing; English.
XX
The invention relates to identifying target proteins
(CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
aligning and comparing nucleic acid or amino acid sequences from plant
with nucleic acid or amino acid sequences from non-plant organisms using
suitable search parameters, where plant sequences having an E-value
greater by a factor of 3 than the E-value of most similar non-plant
```





PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	10-AUG-1999;	99US-0147935.
PR	11-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 11.9%; Score 171.5; DB 21; Length 277;		
Best Local Similarity 27.0%; Pred. No. 2.8e-08;		
Matches 53; Conservative 32; Mismatches 70; Indels 41; Gaps 9;		
Qy	8	SGAELYTLEEVOYKFEARMKMAA--ASGTSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
Db	55	SGSGFQSNQOEFLYCKAEVQMKLVPCNSAGTVTFYLK-----SGTWDIEDFEFLGNI 108
Qy	65	--NPGSFQSNIIITGKAGAKTSEKHHAVSPAADQAFHTYGLWTPNYVRWTVDGQEVKRT 122
Db	109	SGHPYTLHTNVYT-KGSGDKEOQFHLWEDPTAN--FHTYCITWNQRIIFTVDGPIPIREF 165
Qy	123	EGGQVSNLTGTQGLRF-----NLWSSSESAWVG---QFDESKLPLFOFIN----- 164
Db	166	MNAE-----SRGVFPPTKQPMRLYASLWEAEHWATRGLEKTDWSKAPFTTAYYRNVE 219
Qy	165	---WVKYKYPGQGE 177
Db	220	GCVVVNGKSVCPANSQ 235
Search completed: January 9, 2003, 12:16:18		
Job time : 28.8125 secs		



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:13:37 ; Search time 22.1019 Seconds  
(without alignments)  
2312.009 Million cell updates/sec

Title: US-09-654-652A-1  
Perfect score: 1333  
Sequence: 1 MVSAGFSGAELTYLEVOY.....TRKGQSFNGQYPRDDEPAP 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organellae.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	266.5	20.0	851	16	Q9K7X5
2	248.5	18.6	256	2	Q9APD8
3	247.5	18.6	276	2	Q45648
4	242.5	18.2	214	2	Q93G68
5	242.5	18.2	802	2	Q9S310
6	241.5	18.1	214	2	Q93GE7
7	235.5	17.7	214	2	Q93GE6
8	234.5	17.6	239	2	Q45691
9	232.5	17.4	242	2	Q8RMP0
10	230	17.3	212	2	Q9FDC9
11	230	17.3	237	2	O07856
12	221	16.6	246	16	Q97FD3
13	219	16.4	205	2	Q93UI1
14	206	15.5	289	16	Q8U8N5
15	200	15.0	293	16	Q98C78
16	190.5	14.3	302	16	Q9K7X6

17	189	14.2	245	3	O14412
18	184	13.8	282	10	Q9FI31
19	172	12.9	282	10	O80803
20	171.5	12.9	277	10	Q9MOD1
21	169.5	12.7	282	10	Q9MOD2
22	169	12.7	642	16	Q9WXN1
23	168.5	12.6	277	10	Q94A49
24	167.5	12.6	302	3	O42800
25	167	12.5	646	2	O60039
26	163	12.2	286	10	Q80910
27	161	12.1	284	10	Q9FKL8
28	159	11.9	286	10	O8W4M6
29	159	11.9	287	10	Q9Z5U4
30	157.5	11.8	92	3	Q9HGU1
31	157.5	11.8	94	3	Q9P420
32	156	11.7	292	10	Q9ZRV1
33	154.5	11.6	277	10	Q38907
34	154.5	11.6	284	10	Q9SEB0
35	154.5	11.6	306	2	O51333
36	154	11.6	845	2	O9KWF3
37	154	11.6	845	16	O8XNF8
38	150.5	11.3	269	10	Q39148
39	149	11.2	163	2	Q93UI2
40	146.5	11.0	284	10	Q38857
41	145.5	10.9	285	10	Q9FKL9
42	143	10.7	435	2	O68641
43	141.5	10.6	285	10	Q8S902
44	141.5	10.6	289	10	P93669
45	141	10.6	280	10	Q94910

## ALIGNMENTS

### RESULT 1

Q9K7X5 ID Q9K7X5 PRELIMINARY; PRT; 851 AA.  
AC Q9K7X5;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73).  
GN BGLS OR BH3232.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=C-125 / JCM 9153;  
RC MEDLINE=20512582; Pubmed=11058132;  
RX Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR HSSP; P23904; IAJK.  
DR EMBL; AP001518; BAB06951.1; .  
DR InterPro; IPR000757; Glyco\_hydro\_16.  
DR InterPro; IPR001899; Gram\_pos\_anchor.  
DR Pfam; PF00722; Glyco\_hydro\_16; 3.  
DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
DR PRINTS; PR00737; GLHYDRIASE16.  
DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 3.  
KW Hydrolase; Glycosidase; Complete proteome.  
SQ SEQUENCE 851 AA; 98142 MW; C1C4F0F158400285 CRC64;

Query Match 20.08; Score 266.5; DB 16; Length 851;  
Best Local Similarity 29.7%; Pred. No. 1.8e-13;  
Matches 82; Conservative 31; Mismatches 108; Indels 55; Gaps 12;  
QY 4 AKDFSGAELTYLEVOYKGFEARKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLG 63



Db 112 GKDTTKVQFNYYTNGAG---NHEKLDLGFDAANAYHTYAFDWPQNSIKWYVDGQ-LKHT 167

QY 123 EGGQVSNLTGTGGLRNFNLWSSESA-AWVGQFDESKLPFLQFQFNINWVKVK 170

Db 168 APTQIPAAPGK--IMNWLNGTGVDWLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 5

Q9S310 PRELIMINARY; PRT; 802 AA.

AC Q9S310;

DT 01-MAY-2000 (TrenBLrel. 13, Created)

DT 01-DEC-2001 (TrenBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrenBLrel. 21, Last annotation update)

DE Family 11 xylanase /family 16 beta (1,3-1,4) glucanase.

QXND.

OS Ruminococcus flavefaciens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;

OC Clostridiales; Lachnospiraceae; Ruminococcus.

OX NCBI\_TaxID=1265;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=17;

RC Aurilia V., Martin J.C., Munro C.A., Mercer D.K., Flitt H.J.;

RT "Organisation and strain distribution of genes responsible for the

RT utilization of xylans by the rumen cellulolytic bacterium Ruminococcus

RT flavefaciens 17.";

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ132472; CAB51934.1; -

DR HSSP; P23904; LAJK.

DR InterPro; IPR003305; CBM\_CenC.

DR InterPro; IPR001137; GH\_11.

DR InterPro; IPR000757; Glyco\_hydro\_16.

DR Pfam; PF02018; CBM\_4\_9; 2.

DR Pfam; PF00457; Glyco\_hydro\_11; 2.

DR Pfam; PF00722; Glyco\_hydro\_16; 2.

DR PRINTS; PR00911; GLHYDRLASE11.

DR PRINTS; PR00737; GLHYDRLASE16.

DR PROSITE; PS00776; GLYCOSYL\_HYDROL\_F11\_1; 2.

DR PROSITE; PS00777; GLYCOSYL\_HYDROL\_F11\_2; 2.

DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 2.

FT CHAIN 2 802 FAMILY 11 XYLANASE /FAMILY 16 BETA

FT (1,3-1,4) GLUCANASE.

SQ SEQUENCE 802 AA; 89019 MW; F36BC68805FC5274 CRC64;

Query Match 18.2%; Score 242.5; DB 2; Length 802;

Best Local Similarity 34.9%; Pred. No. 1.6e-11;

Matches 58; Conservative 21; Mismatches 74; Indels 13; Gaps 6;

QY 7 FSGAELYTLVEVQYKGFARMKMAASGTVSSMFLYQNGSEIADGRPHVEVDIEVLGKNP 66

Db 634 YSGGEFRTNNFYHYECESMOAMKNDGVSVSFFTYTGPS---DDNPWDEIDIEILGKNT 690

QY 67 GSFQSNLTGKAGAKTSEKHHAVSPAADOAFHTYGLWTPNYVRWTVDGQEVKRTGEGQ 126

Db 691 TVQVQFNYYTNGGKH---EKLYDLGDFDSEAYHTYGFQDPNPYIAWYVDGREVYRA---- 743

QY 127 VSNLTGTGQ-LRFNFWSSESA-AWVGQFDESKLPFLQFQFNINWVKVK 170

Db 744 TQDIPKTPGKIMNNAWPLGTVDLDLKAEN-GETPLTAHYQWYTYNK 738

RESULT 6

Q93GE7 PRELIMINARY; PRT; 214 AA.

AC Q93GE7;

DT 01-DEC-2001 (TrenBLrel. 19, Created)

DT 01-DEC-2001 (TrenBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrenBLrel. 20, Last annotation update)

DE Endo 1-3,1-4-beta-glucanase (Fragment).

OS uncultured bacterium.

OC Bacteria; environmental samples.

OX NCBI\_TaxID=77133;

RN [1]

RP SEQUENCE FROM N.A.

RA Rodriguez V., Mellado R.P.; "

RT "Novel licheneses from soil."

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF254959; AAK50612.2; -

DR InterPro; IPR000757; Glyco\_hydro\_16.

DR Pfam; PF00722; Glyco\_hydro\_16; 1.

DR PRINTS; PR00737; GLHYDRLASE16.

DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; UNKNOWN\_1.

FT NON\_TER 1

SQ SEQUENCE 214 AA; 24175 MW; 3DB192D2F6B4CAA3 CRC64;

Query Match 18.1%; Score 241.5; DB 2; Length 214;

Best Local Similarity 35.5%; Pred. No. 3.3e-12;

Matches 60; Conservative 24; Mismatches 74; Indels 11; Gaps 6;

QY 3 SAKDFSAGELYTLVEVQYKGFARMKMAASGTVSSMFLYQNGSEIADGRPHVEVDIEVL 62

Db 55 SYNKFDCGGENRSVQTYGYGLYEVRMKPAKNTGIVSFFTYTGPT---GTPWDEIDIEFL 111

QY 63 GKNPGSFQSNLTGKAGAKTSEKHHAVSPAADOAFHTYGLWTPNYVRWTVDGQEVKRT 122

Db 112 GKDTTKVQFNYYTNGAG---NHEKLDLGFDAANAYHTYAFDWPQNSIKWYVDGQ-LKHT 167

QY 123 EGGQVSNLTGTGGLRNFNLWSSESA-AWVGQFDESKLPFLQFQFNINWVKVK 170

Db 168 APTQIPAAPGK--IMNWLNGTGVDWLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 7

Q93GE6 PRELIMINARY; PRT; 214 AA.

AC Q93GE6;

DT 01-DEC-2001 (TrenBLrel. 19, Created)

DT 01-DEC-2001 (TrenBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrenBLrel. 20, Last annotation update)

DE Endo 1-3,1-4-beta-glucanase (Fragment).

OS uncultured bacterium.

OC Bacteria; environmental samples.

OX NCBI\_TaxID=77133;

RN [1]

RP SEQUENCE FROM N.A.

RA Rodriguez V., Mellado R.P.; "

RT "Novel licheneses from soil."

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF254961; AAK50614.2; -

DR InterPro; IPR000757; Glyco\_hydro\_16.

DR Pfam; PF00722; Glyco\_hydro\_16; 1.

DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; UNKNOWN\_1.

FT NON\_TER 1

SQ SEQUENCE 214 AA; 24131 MW; 1ESD8BFA4300EBA5 CRC64;

Query Match 17.7%; Score 235.5; DB 2; Length 214;

Best Local Similarity 35.5%; Pred. No. 1e-11;

Matches 60; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

QY 3 SAKDFSAGELYTLVEVQYKGFARMKMAASGTVSSMFLYQNGSEIADGRPHVEVDIEVL 62

Db 55 SYNKFDCGGENRSVQTYGYGLYEVRMKPAKNTGIVSFFTYTGPT---GTPWDEIDIEFL 111

QY 63 GKNPGSFQSNLTGKAGAKTSEKHHAVSPAADOAFHTYGLWTPNYVRWTVDGQEVKRT 122

Db 112 GKDTTKVQFNYYTNGAG---NHEKLDLGFDAANAYHTYAFDWPQNSIKWYVDGQ-LKHT 167

QY 123 EGGQVSNLTGTGGLRNFNLWSSESA-AWVGQFDESKLPFLQFQFNINWVKVK 170

Db 168 APTQIPAAPGK--IMNWLNGTGVDWLGSYNGVN-PLYAHYDWRVRYTK 213

RESULT 8

Q45691 PRELIMINARY; PRT; 239 AA.

ID Q45691



```
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16;
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW SIGNAL.
FT CHAIN
FT CHAIN
FT CHAIN
SQ SEQUENCE 237 AA; 26989 MW; 7DEF5BCE53790470 CRC64;
Query Match 17.3%; Score 230; DB 2; Length 237;
Best Local Similarity 33.1%; Pred. No. 3.3e-11;
Matches 54; Conservative 26; Mismatches 69; Indels 14; Gaps 6;
QY 7 FSGAEALYLEEVQYKGFARKMAAASGVSSMFLYQNGSEIADGRPWVEVDIEVLGNKP 66
DB 81 YTGGEWSKERFGYGLFQVNMKPIKPNQVSSFFTYTGPS---DGTKWDEIDIEFLGKDT 137
QY 67 GSFQSNITGKAGAKTSEKHHAHVSPAADOAFHTYGLWTPNYVWTVDSQGVKRVKTEGGO 126
DB 138 TKVQFNYYT---SGGNIHLYNLGFDASQGFHTYGFQWQADHITWYVDGRAVYTA---- 190
QY 127 VSNLACTQG-LRFNLW--SSSAANVVGQFDESKLPLFQFINWV 166
DB 191 YNNIPSTPCIMMNAWPGTHEVDSWLGAVN-GRTPLYAYYDHI 232
RESULT 12
QY7FD3 PRELIMINARY; PRT; 246 AA.
AC QY7FD3;
DT 01-OCT-2001 (TremBLrel. 18, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Endo-1,3(4)-beta-glucanase family 16.
GN CAC2807.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Giu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007778; AAK80751.1;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PRINTS: PR00737; GLHYDRLASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Complete proteome.
SQ SEQUENCE 246 AA; 27717 MW; C0EBC302678D2FE1 CRC64;
Query Match 16.6%; Score 221; DB 16; Length 246;
Best Local Similarity 33.7%; Pred. No. 1.9e-10;
Matches 58; Conservative 19; Mismatches 73; Indels 22; Gaps 8;
QY 7 FSGAEALYLEEVQYKGFARKMAAASGVSSMFLYQNGSEIADGRPWVEVDIEVLGNKP 66
DB 91 YAGGEYRNNRYGYGLYRVSMKPAKHIGVDSFFSYTGPS---DNNPWEIDIEFLGKDT 147
QY 67 GSFQSNITGKAGAKTSEKH---HAVSPAADOAFHTYGLWTPNYVWTVDSQGVKRVKTE 123
DB 148 TEVQFNYYTNGV-----KHEYLKLGFDASKGFHTYGIWEQNYIAWLDGKVEYRA- 200
QY 124 GGOVSNLACTQG-LRFNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYP 173
DB 201 ---TSNIPHTPGKVMNMLWPGIGVDSWLGAYD-GVTPVKAYYNWA---YNP 245
```

```
RESULT 13
QY3UI1 PRELIMINARY; PRT; 205 AA.
AC QY3UI1;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Endo 1-3,1,4-beta-glucanase (Fragment).
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez V., Mellado R.P.;
RT "Novel lichenases from soil."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF254960; AAK50613.1;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 205
SQ SEQUENCE 205 AA; 22890 MW; E7CADAAFE02A2F51 CRC64;
Query Match 16.4%; Score 219; DB 2; Length 205;
Best Local Similarity 36.2%; Pred. No. 2.2e-10;
Matches 55; Conservative 20; Mismatches 67; Indels 10; Gaps 5;
QY 3 SAKDFSGAEALYLEEVQYKGFARKMAAASGVSSMFLYQNGSEIADGRPWVEVDIEVL 62
DB 55 SYNKFDGCGENRSVQYGYGLYVRMKPAKNTGIVSSFFTYTGTE---GTPWDEIDIESL 111
QY 63 GKNPGSFSQSNITGKAGAKTSEKHHAHVSPAADOAFHTYGLWTPNYVWTVDSQGVKRT 122
DB 112 GKDTTKVGFNYTNGAG---NHEKLDLGFDAANAYHAFDQPNKSIKYYVDGO-LKHT 167
QY 123 EGGOVSNLGTQGLRFLNWSSESA-AWVGQFD 153
DB 168 ATTQVPAAPGK---IMNLLWNGTGVDWLGSYN 197
QY3UI1 PRELIMINARY; PRT; 289 AA.
AC QY3UI1;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Endo-1,3-1,4-beta-glycanase.
GN EXOK OR ATU4055 OR AGR.L1600.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58."
RL Science 294:2317-2323(2001).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009336; AAL44856.1; ALT_INIT.
DR EMBL; AE008279; AAK89373.1; -.
KW Complete proteome.
SQ SEQUENCE 289 AA; 32368 MW; 0384C6F83320EAC9 CRC64;

Query Match 15.5%; Score 206; DB 16; Length 289;
Best Local Similarity 35.1%; Pred. No. 4.le-09;
Matches 54; Conservative 26; Mismatches 56; Indels 18; Gaps 6;

Qy 5 KDFSGAELYTLVEOYKGFARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
Db 115 RFNACGEIQTGRYGRYGYEARMKRAATGSGLSNAFFYIGPT---DKPHDEIDFEVLGK 171

Qy 65 NPGSFQSN-IITGKAGAAKTSEKHHAVSPAADQAFHTYGLWETPNYVRWTVDGOEVRKTE 123
Db 172 NTGKVQLAQYIAAKGG---NEKLVPEGGADAGFNDYAFVWEQRLRYVYNGKLHV--- 224

Qy 124 GQVSNLTP---GTQGLRFNLWSSESAA-WYVQGF 152
Db 225 --EVTDETQIPQNAQKIFPSLWGTDTLKDWMGKF 256

RESULT 15
Q98C78 PRELIMINARY; PRT; 293 AA.
AC Q98C78;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Endo-1,3-1,4-beta-glycanase, ExoK.
GN MLR5264.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_taxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51743.1; -.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR PRINTS; PR00737; GLHYDRLASE16.
KW Complete proteome.
SQ SEQUENCE 293 AA; 32347 MW; 93BBA1672176ED3C CRC64;

Query Match 15.0%; Score 200; DB 16; Length 293;
Best Local Similarity 33.8%; Pred. No. 1.3e-08;
Matches 51; Conservative 27; Mismatches 61; Indels 12; Gaps 6;

Qy 5 KDFSGAELYTLVEOYKGFARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGK 64
Db 112 REFACGEIQTGRYGRYGYEARLKTGTGSGLSNAFFYIGPS---DKQPWDEIDFEILTK 169
```

```
Qy 65 NPGSFQSN-IITGKAGAAKTSEKHHAVSPAADQAFHTYGLWETPNYVRWTVDGOEVRK-T 122
Db 169 DTSKVQVNAVYIDGKGKNEKLVE----VPGGTDKAFNDYAFVWEKDSLRLWYVNGOLVNTIT 224

Qy 123 EGGQVSNLTGTQGLRFNLWSSES-AAWVQGF 152
Db 225 DPAKLP--SHAQKIFFSLWGSSETMKGMWGA 253

Search completed: January 9, 2003, 12:18:10
Job time : 24.1019 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:04:12 ; Search time 10.9074 Seconds  
(without alignments)  
1613.076 Million cell updates/sec

Title: US-09-654-652A-1  
Perfect score: 1333  
Sequence: 1 MVSADFGAELTYLEVOY.....TRKQESPNQGVPRDEPAP 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 263044 seqs, 70945482 residues

Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2.6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2.6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2.6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2.6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2.6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2.6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2.6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116.5	8.7	673	US-10-294-561-3	Sequence 3, Appli
2	104	7.8	1510	US-09-724-676-83340	Sequence 83340, A
3	104	7.8	1510	US-09-724-676A-83340	Sequence 83340, A
4	103.5	7.8	1554	US-09-724-676-83341	Sequence 83341, A
5	103.5	7.8	1554	US-09-724-676A-83341	Sequence 83341, A
6	99	7.4	1558	US-09-724-676-83343	Sequence 83343, A
7	99	7.4	1558	US-09-724-676A-83343	Sequence 83343, A
8	98.5	7.4	1602	US-09-724-676-83344	Sequence 83344, A
9	98.5	7.4	1602	US-09-724-676A-83344	Sequence 83344, A
10	97.5	7.3	1629	US-09-724-676-83342	Sequence 83342, A
11	97.5	7.3	1629	US-09-724-676A-83342	Sequence 83342, A
12	86.5	6.5	444	US-10-092-411A-4346	Sequence 4346, Ap
13	83.5	6.3	486	US-09-724-676-60984	Sequence 60984, A
14	83.5	6.3	486	US-09-724-676A-60984	Sequence 60984, A
15	83.5	6.3	638	US-09-917-384A-9	Sequence 9, Appli
16	83.5	6.3	640	US-09-917-384A-5	Sequence 5, Appli
17	83.5	6.3	1121	US-09-917-384A-1	Sequence 1, Appli
18	81	6.1	246	PCT-US02-36496-1920	Sequence 1920, Ap
19	81	6.1	246	US-10-293-418-1920	Sequence 1920, Ap
20	80.5	6.0	246	PCT-US02-36496-1264	Sequence 1264, Ap
21	80.5	6.0	246	US-10-293-418-1264	Sequence 1264, Ap
22	80.5	6.0	250	PCT-US02-36496-1280	Sequence 1280, Ap
23	80.5	6.0	250	US-10-293-418-1280	Sequence 1280, Ap
24	78.5	5.9	248	PCT-US02-36496-861	Sequence 861, App
25	78.5	5.9	248	PCT-US02-36496-1037	Sequence 1037, Ap
26	78.5	5.9	248	PCT-US02-36496-1248	Sequence 1248, Ap

27	78.5	5.9	248	1	PCT-US02-36496-1254	Sequence 1254, Ap
28	78.5	5.9	248	1	PCT-US02-36496-1267	Sequence 1267, Ap
29	78.5	5.9	248	6	US-10-293-418-861	Sequence 861, App
30	78.5	5.9	248	6	US-10-293-418-1037	Sequence 1037, Ap
31	78.5	5.9	248	6	US-10-293-418-1248	Sequence 1248, Ap
32	78.5	5.9	248	6	US-10-293-418-1254	Sequence 1254, Ap
33	78.5	5.9	248	6	US-10-293-418-1267	Sequence 1267, Ap
34	78.5	5.9	248	6	US-10-293-418-1274	Sequence 1274, App
35	78	5.9	1125	6	US-10-276-781-1219	Sequence 1219, Ap
36	77.5	5.8	247	1	PCT-US02-36496-1257	Sequence 1257, Ap
37	77.5	5.8	247	6	US-10-293-418-1257	Sequence 1257, Ap
38	77.5	5.8	248	1	PCT-US02-36496-1031	Sequence 1031, Ap
39	77.5	5.8	248	1	PCT-US02-36496-1040	Sequence 1040, Ap
40	77.5	5.8	248	1	PCT-US02-36496-1237	Sequence 1237, Ap
41	77.5	5.8	248	1	PCT-US02-36496-1246	Sequence 1246, Ap
42	77.5	5.8	248	1	PCT-US02-36496-1251	Sequence 1251, Ap
43	77.5	5.8	248	6	US-10-293-418-1031	Sequence 1031, Ap
44	77.5	5.8	248	6	US-10-293-418-1040	Sequence 1040, Ap
45	77.5	5.8	248	6	US-10-293-418-1237	Sequence 1237, Ap

ALIGNMENTS

RESULT 1  
US-10-294-561-3  
; Sequence 3, Application US/10294561  
; GENERAL INFORMATION:  
; APPLICANT: Seikagaku Corporation  
; TITLE OF INVENTION: (1-3)-Beta-D-Glucan Binding Domain Protein, Measuring Method U  
; FILE REFERENCE: Q72918  
; CURRENT APPLICATION NUMBER: US/10/294,561  
; CURRENT FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: JP 2001-351943  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 673  
; TYPE: PRT  
; ORGANISM: Tachypleus tridentatus  
US-10-294-561-3

Query Match	8.7%	Score 116.5	DB 6	Length 673
Best Local Similarity	22.8%	Pred. No. 0.014		
Matches	61	Conservative	43	Mismatches 91
				Indels 73
				Gaps 16
QY	1	MVSAK--DFSGAELTYLEEV-----QYGFARKMAAASGTVSSMFLYONGSEIAD	50	
Db	73	VITAKREDYDGRK-YTSARLKTQFDKSKWKYKTEAKMAIPSRG-VVWMF-----	120	
QY	51	GRPWVEVDIEVLGKNPGSQSNITICK-AGAOKTSEKHIAVSPAADQA-----	97	
Db	121	---WMSGDNTNYYRWPSSEIDFIEHRTNNEKVRGTIHWSTPDGAHAHNSNTNGID	177	
QY	98	FHTYGLWPNYRWVVDGOEV--RKTEGGQVSNLTGTLRENLSSESAA----WVGQ	151	
Db	178	YHLYSEVWNSIVKWFVNGQYFEVKIQGG---VNGKSAFRKNKVFVILNMAIGGNPG-	232	
QY	152	FD--ESKLPLOFIMNWKVKYKTPGQEGG-----SDFTLDWTDNFTFDG	195	
Db	233	FDVADEAFPAKMYIDVRYVQDASTSSPVGDTSLDGYFVQNRHSELYLDVTDASNE-DG	291	
QY	196	SRWKGDTWTFDGNR-----VLDTKNIY	218	
Db	292	AFLOQ--WSYSGNENQOQDFEHLNNVY	317	
RESULT 2				
US-09-724-676-83340				
; Sequence 83340, Application US/09724676				
; GENERAL INFORMATION:				



Query Match	7.4%;	Score 99;	DB 5;	Length 1558;
Best Local Similarity	18.7%;	Pred. No. 1.8;		
Matches 59;	Conservative 38;	Mismatches 100;	Indels 118;	Gaps 11;
Qy 12	LYTLEEVOY - GKFEARKMAAASGTVSSMFLYONGSEIADGRPMWVEDIVELGNKPGSFQ 70			
Db 1210	LTTRAFRYEARFOOKLLEYTDSNNIASLFL-----TANRWLEVRMEYIG----- 1255			
Qy 71	SNITGTAGAGAKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTDGGEVRKTEGCGQVSNL 130			
Db 1256	ACVVLIAAATVSISNLSHRELSAGLVGLGLTYAL-MVSNLYNMWYRNLMADMELOLGAVKRI 1314			
Qy 131	TGT-----OGL-----RFLNWSSESAAWVG----- 150			
Db 1315	HGLLKTEAESYEGLLGBRLRERGGEESKEECVWVGHHGAWGWTGFGYSCGPCLVLSPA 1374			
Qy 151	-----QDESKLPLEFQFINWVKVYKTPGO-----GE 177			
Db 1375	RPPAPSLIPKNWPQGGKIQIONLSVRYDSSLLKPLKHNAL---ISPGOKIGICGRGTS 1430			
Qy 178	GGSDFTLDWTDNFDFTGDSRWGKGDWTF-DGNR-----VDL 212			
Db 1431	GKSSFSLAFFRMVDTFEGHIIITEGGENFSQGROLFCLARAFVVRKTSIFIMDEATASIDM 1490			
Qy 213	TDKNIYSRDGMLILA 227			
Db 1491	ATENILOKVVMTAF 1505			
RESULT 8				

```

DB 1491 ATENLLQKVVMTAFA 1505

RESULT 8
US-09-724-674-83344
; Sequence 83344, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83344
; LENGTH: 1602
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-83344

Query Match 7.4%; Score 98.5; DB 5; Length 1602;
Best Local Similarity 19.5%; Pred. No. 2.1;
Matches 51; Conservative 33; Mismatches 85; Indels 93; Gaps 9;

Qy 12 LYTLEEVOY-GKFEARKMMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGNKPGSFQ 70
Db 1210 LTTTRAFYEARFOQKLELYTDSNNIASLFL-----TANRWLEVRMEYIG----- 1255

Qy 71 SNIITGKAQKTSKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVRRKTEGGQVSNL 130
Db 1256 ACVVLLAAVTSISNSLHRELSAGLVGLGLTYAL-MVSNLYLNMWVRNLADMELQLGAVKRI 1314

Qy 131 TGT-----QGL-----RFLWSESAAWVG----- 150
Db 1315 HGLLKTAEYEGLLGERLRGERGEEKECVWYGGHKGAWGGTGTGSCGPCLVLSPA 1374

Qy 151 -----QFDESKLPFLQFTINWVKVYKTPGQ-----GE 177
Db 1375 RPPAPSLIPKNWPDGQKTIQNLVSRVYDSSLKPKVLKHNAL----ISPGQKIGICGRTGS 1430

Qy 178 GGSDFTLDTWTDNFTDFGSRWG 199
Db 1431 GKSSFSLAFFRMVOTFEQLBG 1452

RESULT 9

```



```
Query Match      6.5%; Score 86.5; DB 6; Length 444;
Best Local Similarity 19.8%; Pred. No. 4.7;
Matches 53; Conservative 35; Mismatches 71; Indels 109; Gaps 14;

Qy 2 VSARDFSA---ELYTLEVOYQKF-EARMKMAAASGTV-SSMFLYQNGSEIADG----- 51
Db 232 MEGKNFGAIIITMSESDSIAYESKQRFNVDTGGYILDNVKIYANGKIGDGFSSSET 291
Qy 52 ----RPWVEVDIEVLGKNFG-SFQSNIIITGKAGAKTSEKHHAVSPAADQAFHTYGLEWT 106
Db 292 YGPDYDDEVIVHAGSETEGTFKSNVNVASASEK----- 327
Qy 107 PNYVRWTVDGOEVRKTEGGQVSNLTGTQGLRNLWSSSAAWQGFDESKLPLFQFINWV 166
Db 328 -----DGG-VTDVT-----VKFDEAID--QYVDKK 350
Qy 167 KVKYTPGQEGGSD-----FTLDWTNDFTDGSRWGKGDTWF- 205
Db 351 LDKEYDDSDSDNDSSGGEVTRENVIDKVESYEGHTLD-TDTYTYKEPEKRTGDKGWGFS 409
Qy 206 ----DGN-----RVDLTDKNI--YSROG 222
Db 410 FLDKXGDLAGSYTVDIDGTYTEYDEDG 437

RESULT 13
US-09-724-676-60984
; Sequence 60984, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60984
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-60984

Query Match      6.3%; Score 83.5; DB 5; Length 486;
Best Local Similarity 22.1%; Pred. No. 9.9;
Matches 30; Conservative 21; Mismatches 48; Indels 37; Gaps 4;

Qy 103 LEWTPNYVRWTVDGOEVRKTEGGQVSNL-----TGTQGLRNLWSSSE 144
Db 149 LHWKVKYLKAILRMKQLEDHEAFETSSLIHGSARVYALYKRDGLLCTGSDDLAKLWDVS 208
Qy 145 SAAWY-----GQFDESKLPLFQFINWVKYKYPG-----QGEGGSDFTLDWTD 188
Db 209 TGQCVYGIQTHCAAVKFEDEOKLVTSFDNTVACWSSSGARTQHFRGHTGAVFSVDYND 268
Qy 189 NFDTFDGSRWGKGDT 204
Db 269 ELDILVS---GSADFT 281

RESULT 14
US-09-724-676A-60984
; Sequence 60984, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60984
; LENGTH: 486
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-60984

Query Match      6.3%; Score 83.5; DB 5; Length 486;
Best Local Similarity 22.1%; Pred. No. 9.9;
Matches 30; Conservative 21; Mismatches 48; Indels 37; Gaps 4;

Qy 103 LEWTPNYVRWTVDGOEVRKTEGGQVSNL-----TGTQGLRNLWSSSE 144
Db 149 LHWKVKYLKAILRMKQLEDHEAFETSSLIHGSARVYALYKRDGLLCTGSDDLAKLWDVS 208
Qy 145 SAAWY-----GQFDESKLPLFQFINWVKYKYPG-----QGEGGSDFTLDWTD 188
Db 209 TGQCVYGIQTHCAAVKFEDEOKLVTSFDNTVACWSSSGARTQHFRGHTGAVFSVDYND 268
Qy 189 NFDTFDGSRWGKGDT 204
Db 269 ELDILVS---GSADFT 281

RESULT 15
US-09-917-384A-9
; Sequence 9, Application US/09917384A
; GENERAL INFORMATION:
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: DING, SHI-YOU
; APPLICANT: MCCARTER, SUZANNE
; APPLICANT: HIMMEL, MICHAEL E.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: VINZANT, TODD B.
; TITLE OF INVENTION: THERMAL TOLERANT EXOGLUCANASE FROM ACIDOTHERMUS
; FILE REFERENCE: NREL 01-38
; CURRENT APPLICATION NUMBER: US/09/917.384A
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Acidothermus cellulolyticus
US-09-917-384A-9

Query Match      6.3%; Score 83.5; DB 5; Length 638;
Best Local Similarity 22.2%; Pred. No. 14;
Matches 36; Conservative 21; Mismatches 56; Indels 49; Gaps 7;

Qy 79 GAQKTSEKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGOEVRKTEGGQVSNLTGTQGLRF 138
Db 375 GATNSWNGNYGTPPAGDSTFYGMAYDWEVY-----HDPPSNNWFG-----F 416
Qy 139 NLWSESAA---WVQFDESKLPLFQFINWVKYKYPGQEGGS---DFTLDWTNDFT 192
Db 417 QAWSMERYAEYYVTGDPKAKLLDKVAWVK-----PNVTTGASWSIPSNLWSGQPD 471
Qy 193 FDGSRWKGDTFDCNRVDLTDKNIYSRDGMILLALTRKGOE 234
Db 472 WNPSNPG-----TNANLH-----VTITSSGQD 493

Search completed: January 9, 2003, 12:14:11
Job time : 12.9074 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:16 ; Search time 6.31481 Seconds  
(without alignments)  
761.932 Million cell updates/sec

Title: US-09-654-652A-1  
Perfect score: 1333  
Sequence: 1 MVSAKDFSGAELTYLEEVQY.....TRKQESFNGQVPRDEPAP 248

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 113974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	120	9.0	287	10	US-09-734-569-152
2	117	8.8	620	10	US-09-988-200-6
3	113	8.5	545	10	US-09-988-200-8
4	86.5	6.5	1385	10	US-09-738-363-2
5	84.5	6.3	1289	10	US-09-738-363-4
6	84	6.3	714	9	US-09-738-626-6377
7	83.5	6.3	422	12	US-10-042-417-4
8	83.5	6.3	848	10	US-09-841-132-192
9	83.5	6.3	1530	10	US-09-841-132-178
10	83.5	6.3	1531	12	US-10-007-693-98
11	80.5	6.0	467	10	US-09-863-5478-1
12	80	6.0	156	10	US-09-734-569-28
13	78.5	5.9	475	9	US-09-738-626-6052
14	77.5	5.8	231	10	US-09-728-911-2
15	77.5	5.8	231	10	US-09-949-192-6
16	77	5.8	200	10	US-09-734-569-150
17	77	5.8	1233	9	US-09-738-626-4312
18	75.5	5.7	77	10	US-09-925-299-865
19	75.5	5.7	731	9	US-09-738-626-4854

20 75.5 5.7 1375 10 US-09-740-274-4 Sequence 4, Appli  
21 75 245 9 US-09-738-626-4876 Sequence 4876, Ap  
22 75 2167 10 US-09-778-927A-61 Sequence 61, Appl  
23 75 2201 12 US-10-100-912-2 Sequence 2, Appli  
24 74 925 10 US-09-452-380-4 Sequence 4, Appli  
25 74 936 10 US-09-452-380-3 Sequence 3, Appli  
26 73.5 356 10 US-09-289-346A-11 Sequence 11, Appl  
27 73.5 678 10 US-09-801-368-314 Sequence 314, App  
28 73.5 862 10 US-09-833-435A-1 Sequence 1, Appli  
29 73.5 930 10 US-09-815-242-13481 Sequence 13481, A  
30 72.5 707 10 US-09-764-870-313 Sequence 313, App  
31 72 395 9 US-10-094-080-3 Sequence 3, Appli  
32 72 433 9 US-09-964-899-11 Sequence 11, Appl  
33 72 544 11 9 US-09-738-626-5606 Sequence 5606, Ap  
34 72 544 10 US-09-854-122-50 Sequence 50, Appl  
35 72 544 1054 9 US-09-738-626-5852 Sequence 5852, Ap  
36 71.5 406 10 US-09-215-450-26 Sequence 26, Appl  
37 71.5 663 10 US-09-815-242-5350 Sequence 5350, Ap  
38 71.5 663 10 US-09-815-242-12138 Sequence 12138, A  
39 71.5 915 10 US-09-332-226-2 Sequence 2, Appli  
40 71.5 54 1050 10 US-09-866-562-38 Sequence 38, Appl  
41 71 345 10 US-09-925-300-1188 Sequence 1188, Ap  
42 71 350 9 US-09-738-626-5115 Sequence 5115, Ap  
43 71 420 10 US-09-905-176-25 Sequence 25, Appl  
44 71 471 10 US-09-765-272-158 Sequence 158, App  
45 71 486 10 US-09-815-242-13455 Sequence 13455, A

#### ALIGNMENTS

##### RESULT 1

US-09-734-569-152  
; Sequence 152, Application US/09734569  
; Patent No. US20020064816A1  
; GENERAL INFORMATION:  
; APPLICANT: Leirchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehrhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Cirpus, Petra  
; APPLICANT: Bischoff, Friedrich  
; APPLICANT: Frank, Markus  
; APPLICANT: Freund, Annette  
; APPLICANT: Duwenig, Elke  
; APPLICANT: Schmidt, Ralf-Michael  
; APPLICANT: Reski, Ralf  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involv  
; FILE REFERENCE: BASF-NAE-1332-99-US  
; CURRENT APPLICATION NUMBER: US/09/734,569  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 60/171,101  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 181  
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1  
; SEQ ID NO 152  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Physcomitrella patens  
US-09-734-569-152

Query Match 9.0%; Score 120; DB 10; Length 287;

Best Local Similarity 23.0%; Pred. No. 0.00037;

Matches 49; Conservative 30; Mismatches 70; Indels 64; Gaps 11;

QY 20 YGKFEARKMAA--ASGTVSSMFLYONGSEIADGRVPEWVIEVLGKNFGS---FOSNII 74

Db 69 YVDISAYIKMPPFDSAGTITFTFMSQQDQ-----HYELDMFLGNTSGQPLLHTNVF 122

QY 75 TKKAGAKTSEKHHAVSPAADAQAFHTYGLWTFNYVWTVGQEVKRTGGQVSNLTGT- 133

Db 123 VDCVGGRG-QQMYLGFDPDSAD--FHYFRFRWSKDMVVFVDNKPVR-----VFKNLEGTV 174









```

RESULT 10
US-10-007-693-98
; Sequence 98, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121-515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 98
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D

```

```

RESULT 11
US-09-863-547B-1
; Sequence 1, Application US/09863547B
; Patent No. US20020128166A1
; GENERAL INFORMATION:
; APPLICANT: Henkel KgaA
; APPLICANT: Hermanus, Lentus B. M.
; APPLICANT: Van Beckhoven, Rudolf F. W. C.
; APPLICANT: Maurer, Karl-Heinz
; APPLICANT: Kottwitz, Beatrix
; APPLICANT: Weiss, Albrecht
; APPLICANT: Van Solingen, Pieter
; TITLE OF INVENTION: Detergents Comprising Cellulases
; FILE REFERENCE: H 1920 A
; CURRENT APPLICATION NUMBER: US/09/863,547B
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 08/945,574
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: PCT/EP96/01755
; PRIOR FILING DATE: 1996-04-26
; PRIOR APPLICATION NUMBER: US 08/614,115
; PRIOR FILING DATE: 1996-03-12
; PRIOR APPLICATION NUMBER: EP 95201115.3
; PRIOR FILING DATE: 1995-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 467
; TYPE: PR1
; ORGANISM: Bacillus sp. CBS 670.93
; NC-09-863-547B-1

```

[illegible]

```

Db 278 MDERNL---SWANWSLTHKDESSAALPGANPTGG-----WTE 312

RESULT 12
US-09-734-569-28
; Sequence 28, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734,569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect 6.1
; SEQ ID NO 28
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-28

Query Match 6.0%; Score 80; DB 10; Length 156;
Best Local Similarity 25.0%; Pred. No. 1.2;
Matches 24; Conservative 19; Mismatches 39; Indels 14; Gaps 5;

QY 20 YGKFEARMKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGS---FQSNII 74
Db 69 YVDISAIYKMPFDSAGVTVTYSQGDQ-----HYELDMFELGNTSGQPFLLHTNVF 122

QY 75 TGKGAQAQKTEKHHAVSPAADQAFHTYGLEWTPNVY 110
Db 123 VDGVGRE-QQMYLGDPSAD--FHYRFRWSKDMV 155

RESULT 13
US-09-738-626-6052
; Sequence 6052, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6052
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6052

Query Match 5.9%; Score 78.5; DB 9; Length 475;
Best Local Similarity 24.5%; Pred. No. 7;
Matches 59; Conservative 25; Mismatches 94; Indels 63; Gaps 14;

QY 14 TLEEYQYQKFEARMKMAAASGTVSSMFLYQNG---SEIADGRPWVEVDIEVLG-----63
Db 142 TGDTTQYQSSAALPFRQWVSGHRLMLGEPGRYISELA--ASWLGV-LAVGGFALLWLR 198

QY 64 -KNPGSFQSNIIITGKGAQAQKTEKHHAVSPAADQ--AFHTY-GLEWTP-----NY 109
Db 199 NKKPGLRKMVTRTGGRGVKTYYRHAALGTVAGLGFVLTFTGLTWSTYAGSNITDLRTQ 258

QY 110 VRWTVDGQEVKRTGEGQVSNLTGTQGLRNLWSSSAAWVGQFD-----ESKLPLFQ 161
Db 259 LNWTOPSVNASLTAAPQVDMHDEHAGHHMHESATSGS--GSIDLVAATAISELRTPL-- 314

QY 162 FINWVKYKYTPGQEGGSDFTLDW--TDNFDTPDGRWKGDMTDFG-----NRVOLT 213
Db 315 -----TITPPAQDG-----LAWTATENRDAY---RFTTDTIAVDGDTGMLTNRLNST 358

QY 214 D 214
Db 359 D 359

RESULT 14
US-09-728-911-2
; Sequence 2, Application US/09728911
; Patent No. US20020012669A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Chen, Zhi
; TITLE OF INVENTION: Human Cytokine Receptor
; FILE REFERENCE: 99-93
; CURRENT APPLICATION NUMBER: US/09/728,911
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/169,049
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/232,219
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/244,610
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-911-2

Query Match 5.8%; Score 77.5; DB 10; Length 231;
Best Local Similarity 25.8%; Pred. No. 3.4;
Matches 34; Conservative 14; Mismatches 41; Indels 43; Gaps 7;

QY 67 GSFQSNIIITGKGAQAQKTEKHHAVSPAADQ-----AFHTYGLEWTPN-----YVRW 112
Db 9 GFLISFLTGVAGTQST---HESLKPQRVQFSRNFHNI-LQWQFGRALTNSSYFYQY 64

QY 113 TVDQEVKRTGEGQVSNLTGTQGLRNLWSSS-----AAWYQDFDESKLPLFQ 161
Db 65 KIYQORQWKNK-----EDCWGQTQELSCDLTSETSDIOEPPYGRVRAASAGSYSE----- 113

QY 162 FINWVKYKYTP 173
Db 162 FINWVKYKYTP 173

```

Db 114 ---WSMTPRFTP 122

RESULT 15

US-09-949-192-6

; Sequence 6, Application US/09949192

; Patent No. US20020142292A1

; GENERAL INFORMATION:

; APPLICANT: Parham, Christi L.

; APPLICANT: Gorman, Daniel L.

; APPLICANT: Kurata, Hirokazu

; APPLICANT: Aral, Naoko

; APPLICANT: Sana, Theodore R.

; APPLICANT: Mattson, Jeanine D.

; APPLICANT: Murphy, Erin E.

; APPLICANT: Savkoor, Chetan

; APPLICANT: Grein, Jeffery

; APPLICANT: Smith, Kathleen M.

; APPLICANT: McClanahan, Terrill K.

; TITLE OF INVENTION: MAMMALIAN GENES; RELATED REAGENTS AND METHODS

; FILE REFERENCE: DX01169K

; CURRENT APPLICATION NUMBER: US/09/949,192

; CURRENT FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: 60/231,267

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 231

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-949-192-6

Query Match 5.8%; Score 77.5; DB 10; Length 231;

Best Local Similarity 25.8%; Pred. No. 3.4;

Matches 34; Conservative 14; Mismatches 41; Indels 43; Gaps 7;

QY 67 GSFQSNITGKAGAKTSEKHAVSPAADQ---AFHTYGLEWTPN-----YVRW 112

Db 9 GFLISFLLGVAGTQST---HESLKPQVQFQSRNFHNI-LWQPCGRALTGNSSVYFVQY 64

QY 113 TVDQGEVRKTEGGQVSNLTGTGGLRFLNWSSES-----AAWVGQFDESKLPLFQ 161

Db 65 KIYGQWQWKNK----EDCWGTQELSCDLTSETSDIQEPPYGRVRAASAGSYSE----- 113

QY 162 FINWVKVVKYTP 173

Db 114 ---WSMTPRFTP 122

Search completed: January 9, 2003, 12:14:39

Job time : 7.31481 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:05:57 ; Search time 25.8333 Seconds  
(without alignments)  
1279.206 Million cell updates/sec

Title: US-09-654-652A-1

Perfect score: 1333

Sequence: 1 MVSAKDFSGAELTYLEEVQY.....TRKQESFNCQVPRDDEPAP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	257.5	19.3	308	AAW93001	B. alkalophilus be
2	246.5	18.5	242	AAW95000	Bacillus subtilis
3	244	18.3	214	AAE07317	Barley recombinant
4	244	18.3	239	AAAR06621	Hybrid (1,3-1,4)-p
5	226	17.0	234	AAAR03775	Thermotable beta-
6	226	17.0	237	AAAR05803	Heat-stable endo-b
7	225	16.9	208	AAAB76858	Bacterial lichenas
8	225	16.9	214	AAAB76859	Bacterial lichenas
9	220.5	16.5	237	AAAR06622	Hybrid (1,3-1,4)-p
10	189	14.2	245	AAW37884	Lichenase protein.

11	184	13.8	282	21	AAAG48467	Arabidopsis thalia
12	184	13.8	282	23	ABB93737	Herbicidally activ
13	172	12.9	282	21	AAAG32464	Arabidopsis thalia
14	172	12.9	282	23	ABB91456	Herbicidally activ
15	171.5	12.9	277	21	AAAG52114	Arabidopsis thalia
16	171.5	12.9	277	23	ABB93169	Herbicidally activ
17	169.5	12.7	269	21	AAAG18644	Arabidopsis thalia
18	169.5	12.7	282	21	AAAG18643	Arabidopsis thalia
19	169.5	12.7	282	23	ABB93168	Herbicidally activ
20	169.5	12.7	298	21	AAAG18642	Arabidopsis thalia
21	163	12.2	286	23	ABB93103	Herbicidally activ
22	161	12.1	284	23	ABB93853	Herbicidally activ
23	159	11.9	287	23	ABB93104	Herbicidally activ
24	154.5	11.6	284	23	ABB93854	Herbicidally activ
25	154	11.6	845	22	AAAB98062	Clostridium perfri
26	150.5	11.3	269	21	AAAG20347	Arabidopsis thalia
27	150.5	11.3	269	21	AAAG20350	Arabidopsis thalia
28	150.5	11.3	269	21	AAAG2653	Arabidopsis thalia
29	150.5	11.3	269	23	ABB93167	Herbicidally activ
30	150.5	11.3	290	21	AAAG20346	Arabidopsis thalia
31	150.5	11.3	297	21	AAAG52652	Arabidopsis thalia
32	150.5	11.3	300	21	AAAG20349	Arabidopsis thalia
33	149.5	11.2	269	21	AAAG20329	Arabidopsis thalia
34	149.5	11.2	299	21	AAAG20328	Arabidopsis thalia
35	149.5	11.2	306	17	AAAR97362	Oerskovia beta-1,3
36	147	11.0	247	18	AAW11593	Protein encoded by
37	147	11.0	247	20	AAV08308	A. thaliana merist
38	145.5	10.9	285	23	ABB93852	Herbicidally activ
39	145	10.9	263	18	AAW29455	Oerskovia xanthine
40	143	10.7	204	21	AAAG48468	Arabidopsis thalia
41	143	10.7	435	18	AAW29456	Oerskovia xanthine
42	141	10.6	261	12	AAAR11599	Beta-1,3-glucanase
43	140.5	10.5	303	18	AAW29457	Oerskovia xanthine
44	140.5	10.5	307	21	AAAG40361	Arabidopsis thalia
45	140.5	10.5	310	21	AAAG40360	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1  
AAW93001  
ID AAW93001 standard; Protein; 308 AA.  
XX  
AC AAW93001;  
XX  
DT 19-MAY-1999 (first entry)  
XX  
DE B. alkalophilus beta-glucanase protein.  
XX  
KW Beta-glucanase; enzyme; glucan; removal; lichenan; membrane;  
XX equipment; food industry; brewing.  
XX  
OS Bacillus alkalophilus.  
XX  
FN DE19732751-A1.  
XX  
PD 04-FEB-1999.  
XX  
PF 30-JUL-1997; 97DE-1032751.  
XX  
PR 30-JUL-1997; 97DE-1032751.  
XX  
PA (HENKEL) HENKEL KGAA.  
XX  
PI Hillen W, Maurer K;  
XX  
DR WPI; 1999-122161/11.  
XX  
N-PSDB; NAX02912.  
XX  
PT New Bacillus alkalophilus beta-glucanase protein and gene - useful  
PT for removing glucan and/or lichenan from membranes in the brewing  
PT industry

XX Claim 6; Page 5; 8pp; German.  
 CC This sequence represents a novel beta-glucanase isolated from *Bacillus*  
 CC *alkalophilus* DSM 9956. This enzyme is useful for removing glucan and/or  
 CC lichenan from membranes and equipment in the food industry, especially  
 CC the brewing industry.  
 XX  
 SQ Sequence 308 AA;  
 Query Match 19.3%; Score 257.5; DB 20; Length 308;  
 Best Local Similarity 36.5%; Pred. No. 4.5e-17;  
 Matches 62; Conservative 22; Mismatches 73; Indels 13; Gaps 6;  
 QY 6 DFGAELTYLEEVQYKGFARMKMAAAGCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKN 65  
 DB 96 EYKAGELRTNQYQYGLFEVNMKPAKSTGTVSSLFY-TGPWDNDNDPWEIDIEFLGKD 154  
 QY 66 PGFSQNIITKAGAKQKTSKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKTEGG 125  
 DB 155 TTRVQFNFTNGVG---NNEHYHELGFDASEFNFYAFEMRPESISWYVNGELVYTA--- 208  
 QY 126 QVSNLTGT-QGLRFLNWSESA-AWVGQFDESKLPLFQFINNVKVKYKTP 173  
 DB 209 -TENIPQTPQKIMMNLWPGIGVDGTGTFDGEDTPVVTEDWV---RYTP 254  
 RESULT 2  
 AAP95000  
 ID AAP95000 standard; Protein; 242 AA.  
 AC AAP95000;  
 XX  
 DT 12-FEB-1997 (first entry)  
 XX  
 DE *Bacillus subtilis* lichenase.  
 XX  
 KW Beer production; fermentation; barley; beta-glucan; hydrolysis;  
 KW lichenase.  
 XX  
 OS *Bacillus subtilis* Y-25.  
 XX  
 PN JP01067181-A.  
 XX  
 PD 13-MAR-1989.  
 XX  
 PF 08-SEP-1987; 87JP-0224615.  
 XX  
 PR 08-SEP-1987; 87JP-0224615.  
 PA (ASAK) ASAHI BREWERIES KK.  
 XX  
 DR WPI; 1989-119863/16.  
 DR N-PSDB; AAP95000.  
 XX  
 PT Recombinant plasmid used in beer prodn. - obtd. by integrating  
 PT lichenase gene derived from *Bacillus subtilis*, into vector  
 XX  
 PS Disclosure: Fig 4; 7pp; Japanese.  
 XX  
 CC The lichenase gene from *Bacillus subtilis* Y-25 is used for  
 CC transforming *Bacillus* hosts so that they show increased lichenase  
 CC expression. The recombinant lichenase enzyme produced by the  
 CC transformants is useful in beer production for decomposing beta-  
 CC glucan from barley.  
 XX  
 SQ Sequence 242 AA;  
 Query Match 18.5%; Score 246.5; DB 10; Length 242;  
 Best Local Similarity 35.5%; Pred. No. 4e-16;  
 Matches 60; Conservative 24; Mismatches 74; Indels 11; Gaps 6;  
 QY 3 SAKDFSGAELTYLEEVQYKGFARMKMAAAGCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62

DB 83 SYNKFDCCGNSRVQTYGYGLYEVNRMKPAKNTGIVSSFFTYTGPT---DGTWDEIDIEFL 139  
 QY 63 GKNPGSFSQNIITKAGAKQKTSKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKRT 122  
 DB 140 GKDTTKVQFNFTNGAG---NHEKIVDLGFDAAAYHTYAFDQWQNSIKWYVDGO-LKHT 195  
 QY 123 EGGQVSNLTGTQGLRFLNWSESA-AWVGQFDESKLPLFQFINNVKVKYK 170  
 DB 196 ATNQIPTTPGK--IMNWLNGTGVDEWLGSYGVN-PLAHYDWDVRYTK 241  
 RESULT 3  
 AAEO7317  
 ID AAEO7317 standard; Protein; 214 AA.  
 AC AAEO7317;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Barley recombinant, thermostable (1,3-1,4)-beta-glucanase enzyme.  
 XX  
 KW Barley; transgenic barley malt; carbohydrate degrading enzyme; chicken;  
 KW (1,3-1,4)-beta-glucanase; nutritional value; foodstuff; thermostable.  
 XX  
 OS Hordeum vulgare.  
 XX  
 PN W0200159141-A2.  
 XX  
 PD 16-AUG-2001.  
 XX  
 PF 09-FEB-2001; 2001WO-US04222.  
 XX  
 PR 10-FEB-2000; 2000US-0181473.  
 PR 09-NOV-2000; 2000US-0247126.  
 XX  
 PA (UNIW) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 PI Von Wettstein D, Huang J, Horvath H;  
 DR WPI; 2001-497082/54.  
 XX  
 PT New barley based foodstuff for animals, i.e. chicken comprising  
 PT recombinant carbohydrate degrading enzyme (1,3-1,4)-beta-glucanase -  
 XX  
 PS Claim 8; Page 37-38; 43pp; English.  
 XX  
 CC The present invention relates to a food stuff comprising barley feed and  
 CC transgenic barley malt where the transgenic barley malt comprises a  
 CC recombinant carbohydrate degrading enzyme comprising a (1,3-1,4)-beta-  
 CC glucanase. The food stuff is useful to feed animal, preferably chickens.  
 CC Barley is cheaper than corn, the principal foodstuff of chickens.  
 CC Chickens cannot efficiently utilise barley as an energy source as they  
 CC do not possess a gut enzyme that depolymerises beta-D-glucan, major  
 CC carbohydrate present in the barley endosperm. The invention provides a  
 CC barley based foodstuff which comprises a recombinant carbohydrate  
 CC degrading enzyme improving the nutritional value of the foodstuff.  
 CC The present sequence is barley recombinant, thermostable  
 CC (1,3-1,4)-beta-glucanase enzyme.  
 XX  
 SQ Sequence 214 AA;  
 Query Match 18.3%; Score 244; DB 22; Length 214;  
 Best Local Similarity 35.5%; Pred. No. 5.9e-16;  
 Matches 61; Conservative 26; Mismatches 69; Indels 16; Gaps 8;  
 QY 3 SAKDFSGAELTYLEEVQYKGFARMKMAAAGCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62  
 DB 55 SYNFDCGNSRVQTYGYGLYEVNRMKPAKNTGIVSSFFTYTGPT---GTPWDEIDIEFL 111  
 QY 63 GKNPGSFSQNIITKAGAKQKTSKHHAVSPAADQAFHTYGLEWTPNYVRWTVDGQEVKRT 122  
 DB 112 GKDTTKVQFNFTNGVGGH---EKVISLGFDAKGFHTYAFDQWQPGYIKWYVDG-VLKHT 167





KW Endo-beta-1,3-1,4-glucanase; barley; brewing.  
 XX  
 OS Bacillus macerans.  
 XX  
 PN DD275704-A.  
 XX  
 PD 31-JAN-1990.  
 XX  
 XX 23-SEP-1988; 88DD-0320082.  
 PF  
 XX 23-SEP-1988; 88DD-0320082.  
 PR  
 XX (DEAK ) AKAD WISSENSCHAFT DDR.  
 PA  
 XX Borriass R, Wobus U, Mendel R-R, Baumlein H;  
 PI  
 XX WPI; 1990-210631/28.  
 DR N-PSDB; AAQ05167.  
 XX  
 XX Prepn. of barley plants expressing heat stable beta-glucanase -  
 PT by transforming cells with appropriate vector then regeneration  
 PT giving seeds useful in brewing without conversion to malt.  
 XX  
 PS Disclosure: ; p; German.  
 XX  
 CC The corresp. gene and the plant promoter region are inserted into an  
 CC expression vector and used to transform barley cells. The transformants  
 CC are used to regenerate barley plants which are useful in brewing. This  
 CC protein accumulates in the ripe (but ungerminated) seeds. The  
 CC proportion of these to malted seeds is increased therefore, without  
 CC an unacceptable increase in viscosity.  
 XX  
 SQ Sequence 237 AA;  
 Query Match 17.0%; Score 226; DB 11; Length 237;  
 Best Local Similarity 35.1%; Pred. No. 4.2e-14;  
 Matches 59; Conservative 23; Mismatches 70; Indels 16; Gaps 8;  
 QY 7 FSGAELTYLLEEVQYKFEARKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66  
 DB 82 FCAEVRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138  
 QY 67 GSFQSNITITKAGAKTSEKHAVSPAADQAFHTYGLEWTPNYVRWTVGQEVKRTGEGQ 126  
 DB 139 TRVQFNYYTNGVGGH---EKVISLGLFDASKGHTYAFDQPGYIKWYVDG-VLKHT---A 191  
 QY 127 VSNLTGTQGLRPNLWSSESA-AWVGQFDESKLPLFQFINWVKVYKYT 172  
 DB 192 TANIPSTPGKIMMNLWNGTGVDDWLGSYNGAN-PLYAEYDWW---KYT 235  
 RESULT 7  
 ABB76858  
 ID ABB76858 standard; Protein; 208 AA.  
 XX  
 AC ABB76858;  
 XX  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Bacterial lichenase #1.  
 XX  
 KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200212511-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 27-JUL-2001; 2001WO-ES00303.  
 XX  
 PR 28-JUL-2000; 2000ES-0001922.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an

PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX  
 PI Rodriguez Munoz V, Perez Mellado R;  
 XX  
 DR WPI; 2002-217195/27.  
 DR N-PSDB; ABL53374.  
 XX  
 PT Producing nucleic acid encoding lichenase, useful for producing enzymes  
 PT for improving filtration in brewing, comprises isolating the nucleic  
 PT acid from soil bacteria  
 XX  
 PS Claim 18; Page 18-19; 27pp; Spanish.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an  
 CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)  
 CC activity. The enzyme is useful in brewing and for degrading  
 CC beta-glucanases that cause problems during filtration.  
 XX  
 SQ Sequence 208 AA;  
 Query Match 16.9%; Score 225; DB 23; Length 208;  
 Best Local Similarity 36.2%; Pred. No. 4.3e-14;  
 Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;  
 QY 3 SAKDFSGAELTYLLEEVQYKFEARKMAAASGTVSSMFLYQNGSEIADGRPWVEVDIEVL 62  
 DB 58 SYNKDCGNSVQTYGYGLYEVKMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 114  
 QY 63 GKNPQSFQSNITITKAGAKTSEKHAVSPAADQAFHTYGLEWTPNYVRWTVGQEVKRT 122  
 DB 115 GKDTTKVQFNYYTNGAG---NHEKLDLGLFDAANAYHTYAFDQWQPNISKWYVDGQ-LKHT 170  
 QY 123 EGGQVSNLTGTQGLRPNLWSSESA-AWVGQFD 153  
 DB 171 ATTQTPAAPGK--IMMNLWNGTGVDDWLGSYN 200  
 RESULT 8  
 ABB76859  
 ID ABB76859 standard; Protein; 214 AA.  
 XX  
 AC ABB76859;  
 XX  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Bacterial lichenase #2.  
 XX  
 KW Soil bacterium; enzyme; lichenase; endo-1,3/1,4-beta-glucanase; brewing.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200212511-A1.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 27-JUL-2001; 2001WO-ES00303.  
 XX  
 PR 28-JUL-2000; 2000ES-0001922.  
 XX  
 PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX  
 PI Rodriguez Munoz V, Perez Mellado R;  
 XX  
 DR WPI; 2002-217195/27.  
 DR N-PSDB; ABL53375.  
 XX  
 PT Producing nucleic acid encoding lichenase, useful for producing enzymes  
 PT for improving filtration in brewing, comprises isolating the nucleic  
 PT acid from soil bacteria  
 XX  
 PS Claim 18; Page 21-22; 27pp; Spanish.  
 XX  
 CC The present sequence is a protein sequence from a soil bacterium for an

CC enzyme with lichenase/endo-1,3/1,4-beta-glucanase (E.C. 3.2.1.73)  
 CC activity. The enzyme is useful in brewing and for degrading  
 CC beta-glucanases that cause problems during filtration.  
 XX  
 SQ Sequence 214 AA;  
 Query Match 16.98; Score 225; DB 23; Length 214;  
 Best Local Similarity 36.26; Pred. No. 4.5e-14;  
 Matches 55; Conservative 21; Mismatches 66; Indels 10; Gaps 5;  
 QY 3 SAKDFSGAELTYLLEVOYCKFEARKMAAASCTVSSMFLYQNGSEIADGRPWVEVDIEVL 62  
 DB 63 SYNFDCGNSVOTYGYGLYEVRMKPAKNTGIVSSFFTYTGPTE---GTPWDEIDIEFL 119  
 QY 63 GKNPGSFQSNITITGKAGAAQKTSEKHHAHVSPAADQAFHTYGLWTPNHYVETVDGQEVKRT 122  
 DB 120 GKDTTKVQFNYYTNGAG---NHEKLADLGFDAAAYHTYAFDQWPNISKWYVDGQ-LKHT 175  
 QY 123 EGGQVSNLTGTQGLRFLNWSSESA-AWVGQFD 153  
 DB 176 ATTOIPAAPGK--IMMNLWNGTGVDDWLGSYN 205

RESULT 9  
 AAR06622  
 ID AAR06622 standard; protein: 237 AA.  
 XX  
 AC AAR06622;  
 DT 09-JAN-1991 (first entry)  
 XX  
 DE Hybrid (1,3-1,4)-pre-beta-glucanase.  
 XX  
 KW Hybrid pre-beta-glucanase; glucans; beer; animal feed; poultry.  
 XX  
 OS Bacillus amyloliquefaciens, Bacillus macerans.  
 XX  
 FT Key Location/Qualifiers  
 FT Domain 1..127  
 FT /label=amino-terminal half of B.macerans beta-glucanase  
 FT 131..234  
 FT /label=carboxyl-terminal half of B.amyloliquefaciens  
 FT beta-glucanase  
 XX  
 PN W09009436-A.  
 XX  
 PD 23-AUG-1990.  
 XX  
 PF 16-FEB-1990; 90WO-DK00044.  
 XX  
 PR 04-AUG-1989; 89DK-0003848.  
 PR 16-FEB-1989; 89DD-0325800.  
 XX  
 PA (CARL-) CARLSBERG A/S.  
 PA (DEAK ) AKAD WISSENSCHAFT DDR.  
 XX  
 PI Borriis R, Hofeweister J, Thomsen KK, Olsen O, Vonwettstein D;  
 XX  
 XX WPI: 1990-275129/36.  
 DR N-PSDB; AAR05833.  
 XX  
 XX New thermostable (1,3-1,4)-beta-glucanase - prepd. using hybrid  
 FT gene obtd. using Bacillus amyloliquefaciens and B.macerans genes  
 XX  
 XX Disclosure; page 28; 84pp; English.  
 XX  
 CC This hybrid protein is encoded by the beta-glucanase-H2 gene.  
 CC Following processing of the signal peptide the mature protein  
 CC is produced, comprising the amino terminus of the B.macerans  
 CC beta-glucanase and the carboxyl-terminal half of the B.amylo-  
 CC liquefaciens beta-glucanase. This hybrid protein is thermo-  
 CC stable and hydrolyses beta-glycosidic linkages in (1,3-1,4)-  
 CC beta-glucans. Reducing sugars are obtd. at high temps. and

CC thus this enzyme can be used in the mfr. of food prods., esp.  
 CC beer and animal feed (e.g. for feeding poultry). See also  
 CC AAR05833.  
 XX  
 SQ Sequence 237 AA;  
 Query Match 16.5%; Score 220.5; DB 11; Length 237;  
 Best Local Similarity 33.9%; Pred. No. 1.5e-13;  
 Matches 56; Conservative 23; Mismatches 75; Indels 11; Gaps 6;  
 QY 7 FSGAELTYLLEVOYCKFEARKMAAASCTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNP 66  
 DB 82 FDCAEYRSTNIYGYGLYEVRMKPAKNTGIVSSFFTYTGP---AHGTQWDEIDIEFLGKDT 138  
 QY 67 GSFQSNITITGKAGAAQKTSEKHHAHVSPAADQAFHTYGLWTPNHYVETVDGQEVKRTGEGQ 126  
 DB 139 TKVQFNYYTNGAG---NHEKEADLGFDAAAYHTYAFDQWPNISKWYVDGQ-LKHTATTQ 194  
 QY 127 VSNLTGTQGLRFLNWSSESA-AWVGQFDESKLPFLQFINWVKVYK 170  
 DB 195 IPAAPGK--IMMNLWNGTGVDDWLGSYNGVN-PIYAHYHWMRYK 236

RESULT 10  
 AAW37884  
 ID AAW37884 standard; Protein: 245 AA.  
 XX  
 AC AAW37884;  
 DT 20-AUG-1998 (first entry)  
 XX  
 DE Lichenase protein.  
 XX  
 KW Lichenase; lica; fungus; enzyme; beta-1,4-glucan bond hydrolysis;  
 KW beta-1,3-linked glucan bond; grain-containing feed; grain treatment;  
 KW nutrient availability; brewing; fermentation.  
 XX  
 OS Orpinomyces sp.  
 XX  
 FT Key Location/Qualifiers  
 FT Peptide 1..29  
 FT /note= "signal peptide"  
 FT Protein 30..246  
 FT /note= "mature lichenase"  
 XX  
 PN W09814595-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 03-OCT-1997; 97WO-US17811.  
 XX  
 PR 04-OCT-1996; 96US-0027882.  
 XX  
 PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
 XX  
 PI Chen H, Li X, Ljungdahl LG;  
 XX  
 XX WPI: 1998-240094/21.  
 DR N-PSDB; AAV29067.  
 XX  
 XX New isolated lichenase protein - is obtained from Orpinomyces PC-2,  
 FT used for treatment of grain to improve feeds or to improve brewing  
 FT and fermentation processes  
 XX  
 XX Claim 1; Page 24-25; 41pp; English.  
 XX  
 CC This sequence is the lichenase (lica) of Orpinomyces sp. strain PC-2  
 CC of the invention. The protein was purified from a fungus or a fungal  
 CC culture or from a recombinant DNA molecule having a fungal lichenase  
 CC coding sequence. The lichenase enzyme hydrolyses the beta-1,4-glucan  
 CC bonds adjacent to beta-1,3-linked glucan bonds, but does not cleave  
 CC beta-1,4-linked glucans. They can be used for the treatment of animal  
 CC grain-containing feeds to improve nutrient availability and for treatment



```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 184; DB 21; Length 282;
Best Local Similarity 27.6%; Pred. No. 7.8e-10;
Matches 59; Conservative 33; Mismatches 98; Indels 24; Gaps 9;

QY 2 VSAKDFSGAELTYLLEEVQYKGFARKMAA--ASGTVSSMFLYQNGSEIADGRPWVEVDI 59
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 54 LSLDKFSGSGFQSHOEFLYKVEVQMKLYPGNSAGTVTTFYLK-----SPGTTWDEIDF 107
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 60 EVLKG---NPGSFQSNITGKAGAOKTSEKHHAVSPAADOAHPHYCLEWTPNVYVRVTVDG 116
|| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 108 EFLGNISGHPTLHTNVTY-KGTGDKQQFHLWFDPTVD--FHTYCIWNQKRVKVAQ 164
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 117 QEVKRTGEGQVSNL--TGTQGLRF--NLWSESAAWVG---QFDESKLPFLQFINWVKVY 169
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 165 IPREFKNSEALGVPPKHPMRLYASLWEAEHWATRGGLGKLEKTDNSKAPFTAFYRNYVD 224
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 170 KYTPQGEGGSDFTLWDNFDFTDGS---RWGK 200
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 225 ACWWSNGKSSCSANSWFTQVLDKFGKRVKVAQ 258
:| :||: :| :|| :||: :||: :||: :||: :||: :||: :||: :||: :||:

RESULT 12
ABB93737
ID ABB93737 standard; Protein; 282 AA.
XX
AC ABB93737;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2948.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
(PARB ) BAYER AG.
PA
PI Tietjen K, Weidler M;
XX
WPI; 2002-269010/31.
XX
Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -
XX
Claim 5; SEQ ID NO 2948; 261pp + Sequence Listing; English.
XX
The invention relates to identifying target proteins
(ABB90790-ABB94016) for herbicidally active compounds, comprising
aligning and comparing nucleic acid or amino acid sequences from plant
with nucleic acid or amino acid sequences from non-plant organisms using
suitable search parameters, where plant sequences having an E-value
greater by a factor of 3 than the E-value of most similar non-plant
sequences are selected. The polypeptides or nucleic acids encoding them
are useful for identifying modulators. The identified modulators are
useful as herbicides.
XX

```









PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
<hr/>	
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.
<hr/>	
Query Match 12.9%; Score 171.5; DB 21; Length 277;	
Best Local Similarity 27.0%; Pred. No. 1.3e-08;	
Matches 53; Conservative 32; Mismatches 70; Indels 41; Gaps 9;	
Qy 8	SGAELYTLEEVOYKFEARMKMAA--ASGTSSSMFLYQNGSEIADGRPWVEVDIEVLGK- 64
Db 55	SGSGFQSNQOEFLYKAEVQMKLVPGNSAGTVTFYLK-----SPGTTWDEIDDFELGNI 108
Qy 65	--NPGSFOSNIITGRAGAOKTSEKHHAVSPAADQAFHTYGLWTPNYVWTVDGOEVKRT 122
Db 109	SGHPYTLHTNVYT-KGSGDKQQFHWFDPDTAN--FHTYCITWNPQRIIFTVDGPIREF 165
Qy 123	EGGOVSNLTGTQGLRF-----NLWSSESAAWVG---QFDESKLPLFQFIN----- 164
Db 166	MNAE-----SRGVFPPTKOPMRLYASLMEAEHWATRGGLERTDWSKAPFTAYRYRNVE 219
Qy 165	---WVKVYKYPGQGE 177
Db 220	GCWVWNGKSVCPANSQ 235
<hr/>	
Search completed: January 9, 2003, 12:16:17	
Job time : 27.8333 secs	



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:04:12 : Search time 11.7431 Seconds  
(without alignments)  
1613.076 Million cell updates/sec

Title: US-09-654-652A-2  
Perfect score: 1439  
Sequence: 1 MYSAKDFSGAELYTLEEVQY.....PNSSVDKLAALAEHHHHH 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 263044 seqs, 70945482 residues  
Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	116.5	8.1	673	US-10-294-561-3	Sequence 3, Appl
2	113	7.9	21	US-10-010-160-68	Sequence 68, Appl
3	104.5	7.3	1510	US-09-724-676A-83340	Sequence 83340, A
4	104.5	7.3	1510	US-09-724-676A-83340	Sequence 83340, A
5	103.5	7.2	1554	US-09-724-676A-83341	Sequence 83341, A
6	103.5	7.2	1554	US-09-724-676A-83341	Sequence 83341, A
7	100.5	7.0	526	US-09-455-294A-10	Sequence 10, Appl
8	99.5	6.9	1558	US-09-724-676A-83343	Sequence 83343, A
9	99.5	6.9	1558	US-09-724-676A-83343	Sequence 83343, A
10	98.5	6.8	1602	US-09-724-676A-83344	Sequence 83344, A
11	98.5	6.8	1602	US-09-724-676A-83344	Sequence 83344, A
12	97.5	6.8	1629	US-09-724-676A-83342	Sequence 83342, A
13	97.5	6.8	1629	US-09-724-676A-83342	Sequence 83342, A
14	97	6.7	481	US-10-141-531-59	Sequence 59, Appl
15	94	6.5	290	US-09-424-705B-4	Sequence 4, Appl
16	92	6.4	381	US-10-141-531-47	Sequence 47, Appl
17	92	6.4	381	US-10-141-531-48	Sequence 48, Appl
18	92	6.4	381	US-10-141-531-49	Sequence 49, Appl
19	92	6.4	381	US-10-141-531-50	Sequence 50, Appl
20	92	6.4	381	US-10-141-531-51	Sequence 51, Appl
21	92	6.4	381	US-10-141-531-52	Sequence 52, Appl
22	92	6.4	381	US-10-141-531-53	Sequence 53, Appl
23	92	6.4	381	US-10-141-531-54	Sequence 54, Appl
24	92	6.4	381	US-10-141-531-56	Sequence 56, Appl
25	92	6.4	381	US-10-141-531-58	Sequence 58, Appl
26	90	6.3	378	US-10-114-104A-122	Sequence 122, App

ALIGNMENTS

RESULT 1  
US-10-294-561-3  
; Sequence 3, Application US/10294561  
; GENERAL INFORMATION:  
; APPLICANT: Seikagaku Corporation  
; TITLE OF INVENTION: (1-3)-Beta-D-Glucan Binding Domain Protein, Measuring Method U  
; FILE REFERENCE: Q72918  
; CURRENT APPLICATION NUMBER: US/10/294,561  
; CURRENT FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: JP 2001-351943  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 673  
; TYPE: PRT  
; ORGANISM: Tachypleus tridentatus  
US-10-294-561-3

Query Match	8.1%	Score 116.5:	DB 6:	Length 673:
Best Local Similarity	22.8%	Pred. No. 0.018:	Mismatches 91:	Indels 73:
Matches 61:	Conservative 43:			Gaps 16:
QY	1	MYSAK--DFSGAELYTLEEV-----QYKFEARMKMAASGTVSSMFLYQNGSEIAD	50	
Db	73	VITAKREDYDGFK-YTSARLKTQFDKSKYKTEAKMAIPSRG-VWVMF-----	120	
QY	51	GRPWVEVDIEVGKNGPSQSNITGK-AGAOKTSKHHAVSPAADQA-----	97	
Db	121	--WMSGDNTNVWPSPSGEIDFIEHRNTNNEKVRGTIHWSTPDGAHAHNRESNTNGID	177	
QY	98	FHTYGLEWTPNVVRYTVDGOEV--RKTEGQVSNLTGTGLRNLWSSESA-----WVGQ	151	
Db	178	YHIYVSNSSIVKVFVNGVNFVQFVKIQGG----VNGKSAFRKNKVFVILNMAIGNWPG-	232	
QY	152	FD--ESKLPLOFINWVKVYKYPGQEGG-----SDFTLDWTDNFDTFDG	195	
Db	233	FDVADEAFPAKWIDYVRVYQDASTSSPVGDTSLDGYFVQNRHSELYLDVTDASNE-DG	291	
QY	196	SRWKGDTWTFDGNR-----VDLTDKNIY	218	
Db	292	AFLQO--WSYSGNQOQDFEHLNNVY	317	
RESULT 2				
US-10-010-160-68				
; Sequence 68, Application US/10010160				
; GENERAL INFORMATION:				

; APPLICANT: Rosey, Everett L.  
 ; APPLICANT: Strugnell, Richard A.  
 ; APPLICANT: Good, Robert T.  
 ; APPLICANT: King, Kendall W.  
 ; TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR  
 ; TITLE OF INVENTION: TREATING INFECTION BY LAWSONIA SPP.  
 ; FILE REFERENCE: DAV110.001AUS  
 ; CURRENT APPLICATION NUMBER: US/10/010.160  
 ; CURRENT FILING DATE: 2001-11-09  
 ; PRIOR APPLICATION NUMBER: AU P1381  
 ; PRIOR FILING DATE: 2000-11-10  
 ; PRIOR APPLICATION NUMBER: US 60/249,596  
 ; PRIOR FILING DATE: 2000-11-17  
 ; NUMBER OF SEQ ID NOS: 68  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 68  
 ; LENGTH: 21  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Peptide  
 US-10-010-160-68

Query Match 7.9%; Score 113; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.00036;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 PNSSSVDKLAAALEHHHHH 267  
 Db 2 PNSSSVDKLAAALEHHHHH 21

RESULT 3  
 US-09-724-676-83340  
 ; Sequence 83340, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 83340  
 ; LENGTH: 1510  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-83340

Query Match 7.3%; Score 104.5; DB 5; Length 1510;  
 Best Local Similarity 18.9%; Pred. No. 0.68;  
 Matches 62; Conservative 47; Mismatches 106; Indels 113; Gaps 12;

Qy 12 LYTLEEVQY-GKFEARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70  
 Db 1210 LTTIRAFRIEAFQOKLLEYTDSNNIASLFL-----TAANRWLEVRMEVIG----- 1255  
 Qy 71 SNIITGKAGAQRTSEKHAVSPAADQAFHTYGLEWTPNVRVTVDSQEVKRTGEGQVSNL 130  
 Db 1256 ACVVLIAAVTSSLSHLRELSAGLVGLGLTYAL-MVSNYLNWVRNLADMELOLQCAVKRI 1314  
 Qy 131 TGTQGLRNLWSSESAAMWG-----QFDESKLPFLFOFINWV 166  
 Db 1315 HG-----LLKTEAESYEGLLAPSLIPKNWPDQKIQIQLNSVRYDSSLKPVKLVHVAL 1367  
 Qy 167 KVKYTPGQ-----GEGGSDFTLDWTDNFDFTDGRSGKCGDWTFF-DGNR----- 209  
 Db 1368 ----ISPCQKIGICGRTSGKSSFLAFFRVMVDTFEGHITTEGGENFSQGOQLFCLARA 1423  
 Qy 210 -----VDLTDKNISYRDM-----LILALTRK 231  
 Db 1424 FVRKTSIFIMDEATASIDMATENILQKVMTAFADRTVVTIAHRVHTILSADLVIVLKR- 1482

Qy 232 QGESFNGQVPRDDEPAPNSSSVDKLAAA 259  
 Db 1483 -----GAILEFDKPEKLLSRKDSVFAS 1504  
 RESULT 4  
 US-09-724-676A-83340  
 ; Sequence 83340, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 83340  
 ; LENGTH: 1510  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676A-83340

Query Match 7.3%; Score 104.5; DB 5; Length 1510;  
 Best Local Similarity 18.9%; Pred. No. 0.68;  
 Matches 62; Conservative 47; Mismatches 106; Indels 113; Gaps 12;

Qy 12 LYTLEEVQY-GKFEARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70  
 Db 1210 LTTIRAFRIEAFQOKLLEYTDSNNIASLFL-----TAANRWLEVRMEVIG----- 1255  
 Qy 71 SNIITGKAGAQRTSEKHAVSPAADQAFHTYGLEWTPNVRVTVDSQEVKRTGEGQVSNL 130  
 Db 1256 ACVVLIAAVTSSLSHLRELSAGLVGLGLTYAL-MVSNYLNWVRNLADMELOLQCAVKRI 1314  
 Qy 131 TGTQGLRNLWSSESAAMWG-----QFDESKLPFLFOFINWV 166  
 Db 1315 HG-----LLKTEAESYEGLLAPSLIPKNWPDQKIQIQLNSVRYDSSLKPVKLVHVAL 1367  
 Qy 167 KVKYTPGQ-----GEGGSDFTLDWTDNFDFTDGRSGKCGDWTFF-DGNR----- 209  
 Db 1368 ----ISPCQKIGICGRTSGKSSFLAFFRVMVDTFEGHITTEGGENFSQGOQLFCLARA 1423  
 Qy 210 -----VDLTDKNISYRDM-----LILALTRK 231  
 Db 1424 FVRKTSIFIMDEATASIDMATENILQKVMTAFADRTVVTIAHRVHTILSADLVIVLKR- 1482

RESULT 5  
 US-09-724-676-83341  
 ; Sequence 83341, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 83341  
 ; LENGTH: 1554  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-83341

Query Match 7.2%; Score 103.5; DB 5; Length 1554;  
 Best Local Similarity 20.8%; Pred. No. 0.88;  
 Matches 46; Conservative 36; Mismatches 80; Indels 59; Gaps 8;

Qy 12 LYTLEEVQY-GKFEARMKMAASGTVSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSFQ 70





FILE REFERENCE: 129181.4 Comugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 83342  
LENGTH: 1629  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-83342

Query Match 6.8%; Score 97.5; DB 5; Length 1629;  
Best Local Similarity 19.4%; Pred. No. 3.3;  
Matches 50; Conservative 33; Mismatches 82; Indels 93; Gaps 9;  
QY 12 LYTLEVOY-GRFEARMKAAAGTSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSQF 70  
DB 1210 LTTIRAFREARFQKLLDYTSNNIASLFL-----TAANRWLEVRMEYIG----- 1255  
QY 71 SNIITGAGAOKTSEKHHAVSPAADAQAFHTYGLEWTPNRYVWTVGQEVKTEGGQVSNL 130  
DB 1256 ACVLLIAAVTSISNLSHRELSAGLVGLGLTYAL-MVSNYLNMMVRNLADMEQLGAVKRI 1314  
QY 131 TGT-----QGL-----RFLWSESAAWVG----- 150  
DB 1315 HGLLKTEAESYEGLLGERLRERGGESEKCEVWVGHGKAWGWTGFGYSCGPCLVLSA 1374  
QY 151 -----QFDESKLPFLQFINWVKVYKTPGQ-----GE 177  
DB 1375 RPPAPSLIPKNWPDGKTOIQNLSRYDSSLPVKLVHYNAL-----ISPCQKIGICGRGTS 1430  
QY 178 GGSDFTLDTWNTDFDQ 195  
DB 1431 GKSSFSLAFFRMVDTFEG 1448

RESULT 13  
US-09-724-676A-83342  
Sequence 83342, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Comugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Comugen  
CURRENT APPLICATION NUMBER: US/09/724,676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 83342  
LENGTH: 1629  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-83342  
Query Match 6.8%; Score 97.5; DB 5; Length 1629;  
Best Local Similarity 19.4%; Pred. No. 3.3;  
Matches 50; Conservative 33; Mismatches 82; Indels 93; Gaps 9;  
QY 12 LYTLEVOY-GRFEARMKAAAGTSSMFLYQNGSEIADGRPWVEVDIEVLGKNPGSQF 70  
DB 1210 LTTIRAFREARFQKLLDYTSNNIASLFL-----TAANRWLEVRMEYIG----- 1255  
QY 71 SNIITGAGAOKTSEKHHAVSPAADAQAFHTYGLEWTPNRYVWTVGQEVKTEGGQVSNL 130  
DB 1256 ACVLLIAAVTSISNLSHRELSAGLVGLGLTYAL-MVSNYLNMMVRNLADMEQLGAVKRI 1314  
QY 131 TGT-----QGL-----RFLWSESAAWVG----- 150  
DB 1315 HGLLKTEAESYEGLLGERLRERGGESEKCEVWVGHGKAWGWTGFGYSCGPCLVLSA 1374  
QY 151 -----QFDESKLPFLQFINWVKVYKTPGQ-----GE 177  
DB 1375 RPPAPSLIPKNWPDGKTOIQNLSRYDSSLPVKLVHYNAL-----ISPCQKIGICGRGTS 1430

QY 178 GGSDFTLDTWNTDFDQ 195  
DB 1431 GKSSFSLAFFRMVDTFEG 1448

RESULT 14  
US-10-141-531-59  
Sequence 59, Application US/10141531  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Daimia, Bipin K.  
APPLICANT: Desjarlais, John R.  
APPLICANT: Heifetz, Peter  
APPLICANT: Lugnbuhl, Peter  
APPLICANT: Muchhal, Umesh  
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity  
FILE REFERENCE: A-71457-2/RFT/RMS/RMK  
CURRENT APPLICATION NUMBER: US/10/141,531  
CURRENT FILING DATE: 2002-05-06  
PRIOR APPLICATION NUMBER: US 60/370,609  
PRIOR FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: US 60/289,029  
PRIOR FILING DATE: 2001-05-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 59  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: synthetic  
US-10-141-531-59

Query Match 6.7%; Score 97; DB 6; Length 481;  
Best Local Similarity 20.7%; Pred. No. 0.71;  
Matches 56; Conservative 30; Mismatches 88; Indels 96; Gaps 10;  
QY 79 GAKTSEKHHAVSPAADAQAFHTYGLEWTPNRYVWTVGQEVKTEGGQVSNLTGQLRF 138  
DB 227 GASKIMQORALSNPKIDVTWNSVVEAYGCDGERDVLGGLKRVNVTGVDVSLK-VSGLUFF 285  
QY 139 NLWSSESAAGVQGFDESKLPFLQFINWVKVYKTPGQEGGSDFTLDWTDNF----- 190  
DB 286 AIGHEPATKFLDGGVDSEAEANFLT-----KY-----GSKVYIIHWDAFCASKIMQO 334  
QY 191 -----DTFDGS-----RWGKDWTFDGNRVLDTKNIYSRD-----GMLILALTRKG 232  
DB 335 RALSNPKIDVTWNSVVEAYGCDGERDVLGGLKRVNVTGVDVSLK-VSGLUFFAIGHEP 391  
QY 233 QESF-NGOVPRDDE-----PAPNSS----- 252  
DB 392 ATKFLDGGVELSDGYYVTKPGTTQTTSVPGVFAAGDVQDKYRQAITAAGTCMAALDAE 451  
QY 253 -----VDKLAALAEHHHHHH 267  
DB 452 HYLQEIIGSOQKSDGVDKLAALAEHHHHHH 481

RESULT 15  
US-09-424-705B-4  
Sequence 4, Application US/09424705B  
GENERAL INFORMATION:  
APPLICANT: LITTLE, MELVYN  
APPLICANT: KIPRIYANOV, SERGEY  
APPLICANT: MOLDENHAUER, GERHARD  
APPLICANT: DEUTSCHES KREBSFORSCHUNGSZENTRUM  
TITLE OF INVENTION: MUTATED OKT3 ANTIBODY  
FILE REFERENCE: 035280047US00  
CURRENT APPLICATION NUMBER: US/09/424,705B  
CURRENT FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/DE98/01409  
PRIOR FILING DATE: 1998-05-22

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-424-705B-4

Query Match      6.5%; Score 94; DB 5; Length 290;
Best Local Similarity 20.4%; Pred. No. 0.67;
Matches 66; Conservative 43; Mismatches 91; Indels 124; Gaps 18;

QY 25 ARMKMAASGTVSSMFLYQNGSEIADGRP-----WVE-----VDI 59
   | : | | : | | : | | : | |
Db 10 AGLLLLAAQPAQAQVLOQSGAELA--RFGASVKMSCRASGYTFTTYTMHWVKQRPGQGL 67
   | : | | : | | : | | : | |
QY 60 EVLGK-NPGSFOSNI---ITGKAGAKTSEKHHAVS-----PAADQAFH-----T 100
   | : | | : | | : | | : | |
Db 68 EWIGYINPSRGYTNYNQKFKDKKATLTDDKSSSTAYMOLSSLTSEDSAVYYCARYYDDH 127
   | : | | : | | : | | : | |
QY 101 YGLEWTPNVRVTVDGQEVKTEGGGV-----SNLTCTQGLR--FNLWSSESAAWVGQF 152
   | : | | : | | : | | : | |
Db 128 YSLDWGQCTTLTVSSAKTTPKLGDDILLTQTPASLAVSLGORATISCKASOSVDYDGD- 186
   | : | | : | | : | | : | |
QY 153 DESKLPFLFOFINWVKVYKYPGQ-----GEG-GSDFTLDW--TD 188
   | : | | : | | : | | : | |
Db 187 -----SYLW---YQIIPGPPKLLIYDASNLVSGIPPRFSGSGSGCTDFTLNHPVE 235
   | : | | : | | : | | : | |
QY 189 NFD--TFDGSRWKGDWTFDGNRVDLTDKNYISRDGMLILALTRKGQESFNGQVPRDDEP 246
   | : | | : | | : | | : | |
Db 236 KVDAATYHCQOSTEDPWFEGGT-----KLEIKRADAA 268
   | : | | : | | : | | : | |
QY 247 APNSSSVDKLAAALE---HHHHH 267
   | : | | : | | : | | : | |
Db 269 A--AGSEQKLISEEDLNSHHHHH 290
   | : | | : | | : | | : | |
```

Search completed: January 9, 2003, 12:14:13  
Job time : 13.7431 secs